

Prof:- A K. SAHA.
NUCLEAR PHYSICS DIVISION.
SAHA INSTITUTE OF NUCLEAR PHYSICS,
92, Acharya Prafulla Chandra Road,
CALCUTTA-9.

DR. AJIT KUMAR SAHA.
125, SOUTHERN AVENUE,
CALCUTTA-22

WAVE MECHANICS

[SUKHRAJ RAI READERSHIP LECTURES, 1939-40]

BY

V. V. NARLIKAR

B.A. (Cantab.), B.Sc., F.R.A.S., F.A.Sc., F.N.I.

Professor of Applied Mathematics and
Head of the Mathematics Department,
Benares Hindu University, Formerly Isaac
Newton Student at Cambridge (1930-32)

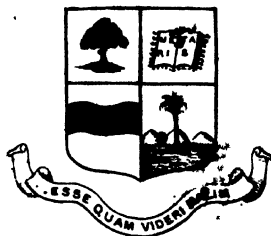
Printed by A. K. SAHA

NUCLEAR PHYSICS DIVISION.

SAHA INSTITUTE OF NUCLEAR PHYSICS.

92, Acharya Pratulla Chandra Road,

CALCUTTA-9.



PATNA UNIVERSITY

1942

Printed by ~~Y.~~ N. KULKARNI, at the Karnatak Printing Press, Chira Bazar, Bombay 2
& published by JAMUNA PRASAD, M.A. (Cal.), B.A., M.Sc. (Cantab.), Registrar,
for the Patna University, Patna.

CONTENTS

CHAPTER		PAGES
	Introduction	v—vii
I.	Classical Mechanics	1—23
	1. Postulates—2. Principles—3. Hamiltonian Systems— 4. Poisson Brackets—5. Conditionally Periodic Systems— 6. Degeneracy—7. Relativistic Expressions—8. The Failure of Classical Mechanics.	
II.	The Quantum Hypothesis	24—31
	1. The Origin of the Quantum Hypothesis—2. Einstein's Relations—3. The Success of the Quantum Hypothesis.	
III.	First Quantum Theories of Matter and Radiation..	32—45
	1. Postulates and Hypotheses of Böhr's Theory—2. Suc- cesses of Böhr's Theory—3. Einstein's Theory of Radiation —4. Failures of the Two Theories.	
IV.	Waves and Particles	46—71
	1. A Fundamental Analogy—2. The Schrödinger Equation —3. The Two Interpretations of ψ —4. Some Collision Phenomena—5. The Kepler Problem.	
V.	Matrices and Waves	72—87
	1. Heisenberg's Matrix Theory—2. The Harmonic Oscillator —3. The Connection between Waves and Matrices—4. Pro- bability Postulates.	
VI.	States and Observables	88—106
	1. A Mathematical Analogy—2. The Calculus of Opera- tors—3. The Postulates and Principles of Dirac's Theory —4. The Uncertainty Principle—5. The Electron Spin.	
VII.	The Atom	107—117
	1. The Problem of n Particles—2. The Energy Levels the Atom—3. The Number of Elements.	
VIII.	Radiation	118—134
	1. The Classical Theory of Radiation—2. The Quantum Theory of Radiation—3. Phenomenological Successes— 4. Peculiar Features and Difficulties—5. The Positive Electron.	

CHAPTER	PAGES
IX. Quantum Theories of Forces	135-156
1. The Nature of Chemical Forces—2. Collectivization of Electrons—3. Some Typical Methods of Calculation— 4. Other Forces of Quantum-Mechanical Nature.	
Name Index	157-158
Subject Index	159-160

INTRODUCTION

The following account of wave mechanics is based on a series of eight lectures delivered by me early in 1940, as Sukhraj Rai Reader in Natural Science of the Patna University. It is too sketchy and succinct to be a textbook or a treatise and, when I undertook to write this, it was for the compelling reason that the Readership lectures *have* to be published. There is no dearth of books and papers on the subject, written in English and American, even by the masters who have invented the theory and applied it to phenomena physical, chemical and astronomical. Most of these authoritative expositions of wave mechanics are meant for the specialist and the advanced research worker and the average M.Sc. student of an Indian University is unable to follow them for lack of adequate mathematical attainment. It is this consideration which has largely influenced the present treatment. A new plan of presentation has been resorted to and in the execution of it lies the merit of the book, if there be any.

From the Poisson brackets to non-commutative relations is a very wide gulf and the theoretical structure that runs across it has been built up part by part by Bohr and Sommerfeld, de Broglie and Schrödinger, Heisenberg and Pauli, Born and Dirac and several others. The many phases through which the quantum theory passed within a few years of this century have left some indelible marks on the final form of the theory. Thus, in the most general theory occur terms and expressions which are described as Hamiltonians, wave-packets and observables. A full understanding of the Dirac phase of the theory is not therefore possible without a knowledge of the theoretical foundations of each earlier phase of quantum physics. Hence it is in the fitness of things that one should start with the shortcomings of the classical theory and climb up, till one reaches the more sophisticated complications of the relativistic quantum mechanics, along the ladder of historical development.

More sophisticated, the complications produce genuine theoretical difficulties. The problems of matter and radiation become inextricably mixed up near the "edge" of the electron mainly on account of the interaction effects and the production of electron pairs. "Inside the electron" is the land of mystery wherein is probably treasured the secret of the electronic self-energy and of the constants c and h . Here are the frontiers of theoretical research in physics and before reaching them we have to pass in review Maxwell's equations, Bohr's selection principles, Einstein's photon-hypothesis, and the quantum theory of radiation, distinguished from the earlier and un-

successful elaboration of Schrödinger's theory to account for the dual nature of light. To make the outline of quantum mechanics complete it is therefore desirable to include an account of radiation from Maxwell to Heisenberg and Dirac. Further reasoning is not needed to justify the inclusion of the theory of radiation here.

The field of application of wave mechanics is very wide. So only a few most prominent applications are touched here. The future of a theory depends upon its inner logic, the present upon its phenomenological success. For this reason, and also to bring out the connection between mathematical abstractions and the observable properties of matter, I have thought it necessary to give a few applications.

The main business of a scientific philosophy is to clarify thought. Words fail to explain the correlation of quantum phenomena which is expressed by symbols. Yet the phenomena themselves are describable by words! It is imperative, therefore, to attempt to harmonise words and symbols in order to produce a clear understanding of the phenomena and their correlation. In the course of our exposition fundamental questions regarding causality, the nature of light, probability, quantum forces, etc. have been examined. But in doing so we have taken special care to keep all metaphysics out. For the treatment of such fundamental questions in relation to science is the subject proper of the philosophy of science; not of metaphysics.

The basic principles of wave mechanics have not been formulated under theological inspiration as the principles of classical mechanics were. This, in itself, is a notable fact proclaiming the advent of a more rational age. The concept of the de Broglie wave was not suggested by a text in the New Testament or the Old. Nothing has inspired the theoretical worker more in this pioneer work than the patient and skilled experimentation in the laboratory. Nothing has affected the course of evolution of physics more effectively than the progressive development of the experimental technique. Observation demands a new theory; the new theory, when it comes, predicts the result of more penetrating observation; the result springs a surprise on the scientist and calls for a more general theory. And so rapidly has the wheel turned in the domain of quantum phenomena that at times only a few months separated the first intimation of the failure of a theory from the dawn of another. These considerations have persuaded me to emphasize, wherever occasion arises, the part played by phenomena in shaping the growth of ideas and theories.

Forbidding, no doubt, is the appearance of the symbolism of wave mechanics. Nevertheless the mathematical operations themselves are extremely simple and straight-forward; the methods have a tremendous power; the logical structure has a beauty of its own. On the formal side of the theory it is purely a calculus of states and observables. In practice one is concerned

with the solution of certain differential equations or with properties of groups. The three varieties of mathematics are introduced in the following pages.

During the months that occupied the writing, whenever I got stuck in my work, I was reminded of the last proposition of Wittgenstein's *TRACTATUS LOGICO-PHILOSOPHICUS*, viz.,

"Whereof one cannot speak, thereof one must be silent".

The critical philosopher would be silent. The speculative philosopher might wax mathematical. There is no other go because "whatever can be said, can be said clearly and whatever cannot be said cannot be said". The reader is well-advised to bear in mind these remarks or, rather, propositions of Wittgenstein while undertaking a serious study of the subject. The author, on his part, has found that when it comes to putting the ideas of modern physics in circulation words are counterfeit coins, symbols are genuine ones. On the other hand, the sense data alone are real and they are describable by words while the rest of the theory, including even the familiar concepts of mass and force, is logical fiction. There is no contradiction here between the last statement and the earlier. One has simply to be alive to the fact that the sensory experiences that are best described by words are best correlated by symbols.

There was a time when the term "wave mechanics" was reserved for the de Broglie-Schrödinger theory giving wave representations of the position and momentum of a particle. Wave mechanics was thus distinguished from "quantum mechanics" or the theory of non-commutative observables and also from the "quantum theory" of Bohr, Sommerfeld and others which was based on classical principles and classically expressed quantum conditions. We use the word "wave mechanics" in the most general sense as Frenkel has done it. It is perfectly legitimate to bring the abstract algebra under wave mechanics since the symbolic operators are linked to experimentally verifiable results only through the eigenstates or wave functions.

One word more before we turn to the march of ideas through the inevitable jungle of symbols. I have to thank the Senate of the Patna University as much for this honour of Readership as for the opportunity to collect thoughts on this subject. I feel particularly grateful to Prin. Kamta Prasad and Prof. D. N. Sen for their unstinted help to make these lectures a success. It is also a pleasure to acknowledge my indebtedness to the Registrar of the University and the Manager of the Karnatak Printing Press, Bombay for their unflinching kindness, courtesy and efficiency in the matter of printing this book.

I CLASSICAL MECHANICS

- | | |
|-------------------------|--|
| 1. Postulates. | 5. Conditionally Periodic Systems. |
| 2. Principles. | 6. Degeneracy. |
| 3. Hamiltonian Systems. | 7. Relativistic Expressions. |
| 4. Poisson Brackets. | 8. The Failure of Classical Mechanics. |

1. POSTULATES

The two basic postulates upon which classical mechanics is built up are the following :

- (A) The physical universe is describable in terms of an absolute Euclidean space and an absolute time.
- (B) The events of the physical universe are governed by the law of causation.

There is a conventional meaning attached to the statements (A) and (B). In the treatment of any problem a fixed frame of reference $O-xyz$ is chosen with respect to which the co-ordinates of material objects are specified. The origin is a fixed point of space and the directions of the axes are fixed with reference to the fixed stars. In the final calculations it is only the relative co-ordinates with respect to some real object that enter and the position of the fictitious fixed origin drops out of the conclusion. The absolute time is measured by the rotation of the earth about its axis relative to the fixed stars. In the second postulate appears the much discussed phrase "the law of causation". We are not concerned here with its most general meaning ; what it means in mechanics is all that matters. Every physical system is characterised by certain parameters ; and, according to the law of causation, the values of the parameters at one instant lead to a unique set of values at any later instant by virtue of the laws of motion. It is implied here that a state of the physical system is completely known when the values of the parameters for that state are obtained. There are three laws of motion given by Newton which are unverifiable and, therefore, of an axiomatic character. We give below the authoritative statement of these laws from Thomson and Tait.¹

"First Law—Every body continues in its state of rest or of uniform motion in a straight line, except in so far as it may be compelled by force to change that state."

Second Law—Change of motion is proportional to force applied, and takes place in the direction of the straight line in which the force acts.

Third Law—To every action there is always an equal and contrary reaction; or, the mutual actions of any two bodies are always equal and oppositely directed."

Consider the motion of a particle of mass m , moving along under the action of a force F . We have the familiar equation,

$$m\ddot{x} = F. \quad (1)$$

When $F = 0$, \dot{x} is constant and the first law is satisfied. In the second law there is the phrase 'change of motion' which really stands for the time-rate of change of $(m\dot{x})$ and so (1) illustrates the second law. If we look upon $(m\ddot{x})$ as the kinematical reaction² to the external force F , then the third law also is found to be illustrated by (1).

Two new concepts are introduced by the three laws; one is that of kinematical mass and the other, that of force. It is necessary to distinguish the kinematical mass from the gravitational mass featuring in Newton's fourth law, the law of universal gravitation. The former is a measure of the kinematical response which a body gives to the force while the latter determines the gravitational pull that a body exercises on another. An exact statement³ of the law of gravitation is appended for reference :

"Every particle of matter in the universe attracts every other particle, with a force whose direction is that of the line joining the two, and whose magnitude is directly as the product of their masses, and inversely as the square of the distance from the each other."

Thought-experiments can be devised to determine the kinematical and gravitational masses of bodies. There is no *a priori* reason why the two should be equal for bodies. Yet they are found to be equal⁴ and that leads to a remarkable simplification of the theory.

Force is a logical fiction invented to explain different states of mind and matter. In the sense in which the term is used in classical mechanics it brings out the inter-connectedness of the components of a system and explains the changes in configuration. The forces of action and reaction are exchange relations between the components of a system or between the system and some external agency. When the external agent is absent or ineffective the system satisfies one or more of the laws of isolation which are the principles of conservation of linear and angular momentum and energy.

The laws of motion are enunciated for structureless bodies or idealised particles. The problem of n bodies is formulated by superposing linearly the equations of equilibrium of all. The procedure was explicitly set out for the first time by d'Alembert. The treatment of complicated problems is considerably simplified by d'Alembert's principle⁵ although its essence is contained in the three laws of motion.

With the weight of accumulated knowledge physical science has burst through both the postulates (A) and (B) ; and wider limits control it on the separate fronts of general relativity and quantum mechanics. The new mathematics of q 's and p 's transcends the orthodox mathematical limitations of the law of causation and of the absolute Euclidean space and time.

Some points of the inner structure of classical mechanics deserve notice. In the first place, the first law of motion cannot be fully reconciled with (A). The state of rest and that of uniform motion are kinematically distinct but dynamically indistinguishable. Secondly, no explanation is given, as we have already stated above, how the kinematical and gravitational masses are equal. The first of these points has led to special relativity while the other constitutes the basis of general relativity. According to the special theory there is no real distinction between the state of rest and that of uniform motion. The general theory goes a step further and liquidates the distinction between the two states and that of acceleration so that there is no need to distinguish the kinematical mass from the gravitational. At the cost of (A) mechanics and gravitation have thus been welded together and a new correlation—the correlation of field laws—between events is set up other than (B).

There are certain other unsatisfactory features, concealed in the postulates of classical mechanics which have been brought to light by the difficulties of quantum mechanics. Newton's laws imply the existence of (1) structureless bodies or particles, (2) distinguishable particles and (3) velocity at an instant for a particle. The assumptions were made for mathematical simplification. The first two were accepted because for a long time no physical situation was known that could challenge them ; the third found favour because the rigorous treatment of limits was unknown in the early days of classical mechanics.

Given a large assembly of similar particles^a which are in rapid motion and whose mean free path is sufficiently small it becomes practically impossible to determine the average velocity of a particular particle since there is nothing to decide whether the particles seen in two different positions at two different instants are different or not. Even if the practical aspects are set aside the theoretical question remains as to whether the laws regarding the different states of the assembly can be deduced from the classical laws of motion. This question has been discussed by Mach⁷ and others and it is found that the required statistical laws cannot be deduced from the classical laws on the assumption that each particle is distinguishable from another. Quantum mechanics is generalised enough to treat such cases.

The difficulty caused by the assumption of structureless bodies persists even in quantum mechanics and the problem of the electronic self-energy and mass still defies solution.

The logical fiction of velocity at an instant is retained even in quantum mechanics but, as we will show later, the uncertainty principle makes up for it

It was Zeno⁸ who first propounded the paradox of the arrow. Suppose that an arrow is in motion at A at some instant ; then it means that after a time it will be at some other point B. But we can also think of the arrow stopping at some intermediate point between A and B. Hence the original assumption that the arrow can go to some point B, no matter how near it is to A, is unwarranted and therefore wrong. From this was concluded that the arrow cannot be in motion at A. The solution of the difficulty is as follows :—

When we see a body moving the impression of motion is cinematographic. What is actually observed is a body in one place at one instant, in another place at another instant, in an intermediate place at an intermediate instant and so on. In an ideal case the situation can be best described by a simple functional relation,

$$s = f(t) \quad (2)$$

between the position co-ordinate s and the time co-ordinate t . To say on observational evidence that s is a continuous function of t is unjustifiable since this cannot be decided by observation. But even if this is permitted for mathematical simplicity we cannot say that ds/dt exists at a particular instant. The average velocity between two neighbouring epochs is observationally real but the limit of the average velocity, ds/dt , is observationally unreal and is a logical fiction. Thus it is possible for bodies to change their places with time so as to produce the false sensory impression that there is a continuous displacement and that a velocity exists at each instant.⁹

So these are some of the considerations which, as will be found later, make the formulation of quantum mechanics on the basis of (A) and (B) impossible.

It is safe to divide the domain of scientific phenomena into three compartments, of which two are already known as microscopic and macroscopic.¹⁰ The former is the domain of quantum laws while the latter presents large-scale phenomena subject to the field laws of general relativity. Between the two is the third compartment, the buffer state, which may as well be given the name 'Mesoscopic'. This is the domain of classical mechanics. As we have already remarked above certain generalizations of the classical theory have led to general relativity ; other generalizations to quantum mechanics. The units of mass, length and time which characterize the three domains may be stated here for comparison.

	M	L	T
Microscopic,	10^{-28} gm	10^{-8} cm	10^{-17} sec
Mesoscopic	1	1	1
Macroscopic	10^{28}	10^{17}	10^8

To produce a theory of mechanics valid for the electron as well as the spiral nebula is the dream of the theoretical physicist. The task is fraught with many difficulties. The classical mechanics fails inside the atom. The quantum mechanics fails "inside the electron".

2. PRINCIPLES

Let us take a dynamical system consisting of n particles and write down the equations of motion of any one of them as

$$m\ddot{x} = X + \dot{X}', \quad m\ddot{y} = Y + \dot{Y}', \quad m\ddot{z} = Z + \dot{Z}', \quad (3)$$

where $X Y Z$ stand for the externally impressed force while $\dot{X}', \dot{Y}', \dot{Z}'$ are the components of the action on m by the surrounding particles. The equations (3) give the equilibrium between the kinematical reaction ($m\ddot{x}, m\ddot{y}, m\ddot{z}$), the external force and the action of the surrounding particles. Similarly, any other particle also is a seat of equilibrium of forces. Hence the joint system of n particles must be in equilibrium in this sense. It follows that the six equations of statical equilibrium of forces can be written for the system. In doing this we need only note that the relevant force at (x, y, z) has the components ($m\ddot{x} - X, m\ddot{y} - Y, m\ddot{z} - Z$) since, in the joint system, all the equal and opposite forces of the nature of $\dot{X}', \dot{Y}', \dot{Z}'$ drop out. The equations of statical equilibrium give

$$\left. \begin{aligned} \Sigma m\ddot{x} &= \Sigma X, \quad \Sigma m\ddot{y} = \Sigma Y, \quad \Sigma m\ddot{z} = \Sigma Z, \\ \Sigma m(y\ddot{z} - z\ddot{y}) &= \Sigma(yZ - zY), \quad \Sigma m(z\ddot{x} - x\ddot{z}) = \Sigma(zX - xZ), \\ \Sigma m(x\ddot{y} - y\ddot{x}) &= \Sigma(xY - yX). \end{aligned} \right\} \quad (4)$$

In the derivation of (4) we have stuck to the terminology of Newton's laws to convince the reader that it is not necessary to evoke d'Alembert's principle which is implicit in them. The linear momentum along Ox , $\Sigma m\dot{x}$ is conserved when $\Sigma X = 0$. The angular momentum about Ox , $\Sigma m(y\dot{z} - z\dot{y})$ is also conserved when the moment of forces $\Sigma(yZ - zY)$ vanishes. The two principles of conservation of linear and angular momentum follow from (4). If we multiply the three equations (3) by $\dot{x}, \dot{y}, \dot{z}$ respectively and add and then combine the n equations so formed for the system,

$$\Sigma m (\dot{x}\ddot{x} + \dot{y}\ddot{y} + \dot{z}\ddot{z}) = \Sigma (X\dot{x} + Y\dot{y} + Z\dot{z}) + \Sigma (\dot{X}'\dot{x} + \dot{Y}'\dot{y} + \dot{Z}'\dot{z}). \quad (5)$$

In the summation,

$$\Sigma (\dot{X}'\dot{x} + \dot{Y}'\dot{y} + \dot{Z}'\dot{z}) = 0, \quad (6)$$

since equal and opposite forces of action and reaction give corresponding to every term of the form $(\dot{X}'\dot{x} + \dot{Y}'\dot{y} + \dot{Z}'\dot{z})$ another, viz., $-(\dot{X}'\dot{x} + \dot{Y}'\dot{y} + \dot{Z}'\dot{z})$.

Therefore,

$$\frac{d}{dt} \Sigma \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \Sigma (X\dot{x} + Y\dot{y} + Z\dot{z}).$$

When the forces are conservative, that is, when a function $V(x, y, z)$ exists such that

$$-\frac{\partial V}{\partial x} = X, \quad -\frac{\partial V}{\partial y} = Y, \quad -\frac{\partial V}{\partial z} = Z,$$

we have the integral of energy,

$$\Sigma \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - \Sigma \int (Xdx + Ydy + Zdz) = \text{const.} \quad (7)$$

or, in the usual notation of the kinetic and potential energy,

$$T + V = \text{constant.} \quad (8)$$

This is one form of the principle of energy.

Suppose that the system of n particles has actually only s degrees of freedom. It may be assumed that s independent co-ordinates q_1, q_2, \dots, q_s specify the configuration at any time t . The co-ordinates of a typical particle of mass m , viz. x, y, z can be taken as functions of the s co-ordinates and the time t . It is clear that

$$\dot{x} = \frac{\partial x}{\partial t} + \Sigma \frac{\partial x}{\partial q_i} \dot{q}_i, \quad \frac{\partial \dot{x}}{\partial \dot{q}_i} = \frac{\partial x}{\partial q_i}, \quad \frac{d}{dt} \left(\frac{\partial x}{\partial q_i} \right) = \frac{\partial \dot{x}}{\partial q_i}. \quad (9)$$

We get from (3)

$$\Sigma m \left(\ddot{x} \frac{\partial x}{\partial q_i} + \ddot{y} \frac{\partial y}{\partial q_i} + \ddot{z} \frac{\partial z}{\partial q_i} \right) = \Sigma \left\{ (X + X') \frac{\partial x}{\partial q_i} + (Y + Y') \frac{\partial y}{\partial q_i} + (Z + Z') \frac{\partial z}{\partial q_i} \right\} \quad (10)$$

$$= Q_i, \quad (11)$$

Q_i being defined in this manner. Now

$$\begin{aligned} \Sigma m \left(\ddot{x} \frac{\partial x}{\partial q_i} + \ddot{y} \frac{\partial y}{\partial q_i} + \ddot{z} \frac{\partial z}{\partial q_i} \right) &= \Sigma m \left(\ddot{x} \frac{\partial \dot{x}}{\partial \dot{q}_i} + \ddot{y} \frac{\partial \dot{y}}{\partial \dot{q}_i} + \ddot{z} \frac{\partial \dot{z}}{\partial \dot{q}_i} \right) \\ &= \frac{d}{dt} \Sigma \frac{\partial}{\partial \dot{q}_i} \left(\frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \right) - \Sigma m \left\{ \dot{x} \frac{d}{dt} \left(\frac{\partial \dot{x}}{\partial \dot{q}_i} \right) + \dot{y} \frac{d}{dt} \left(\frac{\partial \dot{y}}{\partial \dot{q}_i} \right) + \dot{z} \frac{d}{dt} \left(\frac{\partial \dot{z}}{\partial \dot{q}_i} \right) \right\} \\ &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \Sigma m \left\{ \dot{x} \frac{d}{dt} \left(\frac{\partial x}{\partial q_i} \right) + \dot{y} \frac{d}{dt} \left(\frac{\partial y}{\partial q_i} \right) + \dot{z} \frac{d}{dt} \left(\frac{\partial z}{\partial q_i} \right) \right\} \\ &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \Sigma m \left\{ \dot{x} \frac{\partial \dot{x}}{\partial q_i} + \dot{y} \frac{\partial \dot{y}}{\partial q_i} + \dot{z} \frac{\partial \dot{z}}{\partial q_i} \right\} \\ &= \frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}_i} \right) - \frac{\partial T}{\partial q_i}. \end{aligned} \quad (12)$$

There are the s Lagrangian equations of the type,

$$\frac{d}{dt} \left(\frac{\partial T}{\partial \dot{q}} \right) - \frac{\partial T}{\partial q} = Q. \quad (13)$$

When the potential function exists, $Q = \partial V / \partial q$, as considerations of work for virtual displacements δq show immediately.

Hence

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}} \right) - \frac{\partial L}{\partial q} = 0, \quad (14)$$

where $L = T - V$ is called the Lagrangian of the system. The Lagrangian equations¹¹ are valid for more general cases also. They can be deduced by applying the calculus of variation as conditions for the integral $\int_{t_1}^{t_2} L dt$ to be stationary, the terminal values of q and t being assigned. The whole science of dynamics can thus be based on the stationary character of a certain integral,

$$\delta \int_{t_1}^{t_2} L dt = 0; \quad (15)$$

that is, given any arbitrary portion of the trajectory in the q -space if any other path is considered with the same terminal values of q and t the integral over the actual path is either a maximum or a minimum. This is Hamilton's principle. It follows that

$$\begin{aligned} \delta \int_{t_1}^{t_2} L dt &= \int_{t_1}^{t_2} \Sigma \left(\frac{\partial L}{\partial q_i} \delta q_i + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i \right) dt \\ &= \int_{t_1}^{t_2} \Sigma \left(\frac{\partial L}{\partial q_i} \right) \delta q_i dt + \int_{t_1}^{t_2} \Sigma \frac{\partial L}{\partial \dot{q}_i} d \delta q_i \\ &= \left[\Sigma \frac{\partial L}{\partial \dot{q}_i} \delta q_i \right]_{t_1}^{t_2} + \int_{t_1}^{t_2} \Sigma \left\{ \frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) \right\} \delta q_i dt. \end{aligned} \quad (16)$$

The terminal conditions give the first expression as zero. Moreover, since δq_i are arbitrary we have by Hamilton's principle,

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = 0, \quad (17)$$

which are Lagrange's equations. The principle of least action,

$$\delta \int_{t_1}^{t_2} 2 T dt = 0, \quad (18)$$

can be deduced from Hamilton's principle for natural dynamical systems for which T and V do not contain t explicitly, the energy constant being the same for the actual as well as the varied paths. Let

$$\int_{t_1}^{t_2} L dt = W, \quad \int_{t_1}^{t_2} 2 T dt = S$$

and

$$T + V = \alpha_1$$

so that

$$W = S - \alpha_1 (t_2 - t_1).$$

W is Hamilton's function while S is the action function. It is obvious under the conditions of variation that if $\delta W = 0$, $\delta S = 0$ also. When $T = 0$, Hamilton's principle reduces to the principle on which statics is based, viz.,

$$\delta \int V dt = \delta V \int dt = 0 \text{ or } \delta V = 0.$$

3. HAMILTONIAN SYSTEMS

The Hamiltonian equations¹² can also be deduced from the principle,

$$\delta \int_{t_1}^{t_2} L dt = 0. \quad (19)$$

Consider the expression,

$$\Sigma p \dot{q} - L(q, \dot{q}, t) \quad (20)$$

out of which \dot{q} may be eliminated by using the relation,

$$p = \frac{\partial L}{\partial \dot{q}}, \quad (21)$$

defining the generalized momentum p . Let

$$H(p, q, t) \equiv \Sigma p \dot{q} - L(q, \dot{q}, t). \quad (22)$$

Hamilton's principle may now be expressed in the form

$$\delta \int_{t_1}^{t_2} [-H + \Sigma p \dot{q}] dt = 0, \quad (23)$$

that is,

$$\int_{t_1}^{t_2} \left[-\Sigma \frac{\partial H}{\partial q} \delta q - \Sigma \frac{\partial H}{\partial p} \delta p + \Sigma \delta p \dot{q} \right] dt + \left[\Sigma p \delta q \right]_{t_1}^{t_2} - \int_{t_1}^{t_2} \Sigma \delta q \dot{p} dt = 0. \quad (24)$$

It follows that, as terminal values of δq are zero,

$$\frac{\partial H}{\partial q} = -\dot{p}, \quad \frac{\partial H}{\partial p} = \dot{q}. \quad (25)$$

These are the canonical equations of Hamilton.

An important feature of a variational principle is that its statement is independent of the co-ordinate system chosen. If another co-ordinate system is chosen, say, $q' - p'$, we would have in place of (23),

$$\delta \int_{t_1}^{t_2} [\Sigma p' \dot{q}' - H'] dt = 0 \quad (26)$$

and instead of (25)

$$\frac{\partial H'}{\partial q'} = -\dot{p}', \quad \frac{\partial H'}{\partial p'} = \dot{q}'. \quad (27)$$

The equivalence of (23) and (26) for any two terminal conditions suggests that

$$\Sigma p \dot{q} - H(p, q, t) \equiv \Sigma p' \dot{q}' - H'(p', q', t) + \frac{dS}{dt} \quad (28)$$

where S is some function of q, p, q', p' . In particular let

$$S \equiv S(q, q', t). \quad (29)$$

The identity (28) gives

$$p = \frac{\partial S}{\partial q}, \quad p' = -\frac{\partial S}{\partial q'}, \quad H' = H + \frac{\partial S}{\partial t}. \quad (30)$$

The transformation from $p - q$ to $p' - q'$ as given by (30) is called a contact transformation.¹³ The geometrical significance which gives this transforma-

tion its name is that if two hypersurfaces are tangent at a point in the $q-p$ space their transforms are also tangent at the corresponding point in the $\dot{q}-\dot{p}$ space. In the simple case of two dimensions if two curves touch at a point their transforms also touch at the corresponding point in the other space.

There are three other forms of contact transformation possible according to the nature of the function S chosen. It is sufficient to indicate here one other type which is most useful for our applications. Let us have

$$\Sigma p\dot{q} - H \equiv \Sigma \dot{p}\dot{q} - \dot{H} + \frac{d}{dt}[F(q, \dot{p}, t) - \Sigma \dot{p}\dot{q}] \quad (31)$$

so that

$$S \equiv F(q, \dot{p}, t) - \Sigma \dot{p}\dot{q}. \quad (32)$$

Expanding the last term of (31) we get

$$\dot{p} = \frac{\partial F}{\partial q}, \quad \dot{q} = \frac{\partial F}{\partial \dot{p}}, \quad \dot{H} = H + \frac{\partial F}{\partial t} \quad (33)$$

which is another contact transformation between $p-q$ and $\dot{p}-\dot{q}$.

Consider the equation

$$\dot{H} \equiv H\left(\frac{\partial F}{\partial q}, q, t\right) + \frac{\partial F}{\partial t} = 0. \quad (34)$$

This is the well-known Hamilton-Jacobi differential equation. It can be constructed for any problem when the total energy is expressed in the Hamiltonian form, every p being replaced by the corresponding $\partial F/\partial q$. Hamilton's function, $W \equiv \int L dt$, satisfies (34).

The transformation function F may be so chosen that t is not contained explicitly in it. In particular, let the system be conservative, the constant of energy being α_1 . The contact transformation (33) now leaves H invariant, F being obtained from the modified equation,

$$H\left(\frac{\partial F}{\partial q}, q\right) = \alpha_1. \quad (34')$$

We assume that t does not appear explicitly in H the system being conservative. The action function satisfies the Hamilton-Jacobi equation (34).

In either case we proceed further as follows. A complete integral of the equation (34) or (34') will contain n arbitrary constants if there are n q 's, the additive constant of integration being neglected as it plays no part of physical significance: as (33) shows only the partial derivatives of F are significant. Let the n constants of integration be α_r . The function $F(q, \alpha, t)$ gives the transformation, in the first case,

$$\dot{p} = \frac{\partial F}{\partial q}, \quad \dot{\beta} = \frac{\partial F}{\partial \alpha}, \quad \dot{H} = 0 \quad (35)$$

from $p-q$ to $\alpha-\beta$. The Hamiltonian equations of the new system are,

$$\frac{d\alpha}{dt} = 0, \quad \frac{d\beta}{dt} = 0 \quad (36)$$

so that all the α 's and β 's are constants of the motion. By the conditions of contact transformation the co-ordinates and the momenta become expressible as functions of α , β and t and so the problem is fully solved.

In the second case the transformation becomes

$$p = \frac{\partial F}{\partial q}, \quad \beta = \frac{\partial F}{\partial \alpha}, \quad \dot{H} = H = \alpha_1. \quad (35')$$

The Hamiltonian equations are

$$\frac{\partial \beta}{dt} = \frac{\partial \dot{H}}{\partial \alpha}, \quad \dot{\alpha} = \frac{\partial H}{\partial \beta}. \quad (36')$$

As $\dot{H} = \alpha_1$,

$$\dot{\alpha} = 0 = \dot{\beta}_2 = \dot{\beta}_3 = \dots = \dot{\beta}_n, \quad \beta_1 = t + \text{const.}$$

The full solution follows from the equations,

$$\beta = \frac{\partial F}{\partial \alpha},$$

F being determined by solving (34').

As an illustrative example consider the motion of a particle of unit mass under gravity. We have

$$T = \frac{1}{2} (\dot{x}^2 + \dot{y}^2), \quad V = gy$$

and so

$$\frac{\partial T}{\partial \dot{x}} = p_1, \quad \frac{\partial T}{\partial \dot{y}} = p_2.$$

Hence

$$H \equiv T + V = \frac{1}{2} (p_1^2 + p_2^2) + gy.$$

The Hamilton-Jacobi equation now becomes

$$\frac{1}{2} [(\partial F/\partial x)^2 + (\partial F/\partial y)^2] + gy + \frac{\partial F}{\partial t} = 0$$

One solution is

$$F = \alpha_2 x - \alpha_1 t + \int (2\alpha_1 - \alpha_2^2 - 2gy)^{\frac{1}{2}} dy.$$

It may be noticed that α_1 is the constant of the energy integral. We have dropped the additive constant in F since it has no part to play in the dynamical problem.

$$\partial F/\partial \alpha_1 = \beta_1, \quad \partial F/\partial \alpha_2 = \beta_2;$$

$$\text{or} \quad \beta_1 + t = \int \frac{dy}{(2\alpha_1 - \alpha_2^2 - 2gy)^{\frac{1}{2}}}, \quad \beta_2 = x - \int \frac{\alpha_2 dy}{(2\alpha_1 - \alpha_2^2 - 2gy)^{\frac{1}{2}}}.$$

These equations give the time of description and the orbit. To solve this problem one may also proceed with (34').

The special advantage of the Hamiltonian system and the contact transformation is brought out in the perturbation method. Suppose that the equation (34) is too complicated to give an exact solution. It may be possible in such a case to split up H into two parts H_0 and R so that

$$H = R, \quad (37)$$

$$H_0 \left(\frac{\partial F}{\partial q}, q, t \right) + \frac{\partial F}{\partial t} = 0. \quad (38)$$

Using the value of F as given by (38) the co-ordinate system β - α is obtained and the equations (36) are replaced by

$$\frac{\partial R}{\partial \beta} = -\dot{\alpha}, \quad \frac{\partial R}{\partial \alpha} = \dot{\beta} \quad (39)$$

R being expressed in terms of α, β, t .

A peculiarity of the Hamiltonian system of equations deserves notice. Let $\phi(q, p, t)$ be an integral of the system. A variational equation for δq is

$$\frac{d}{dt} \delta q = \sum \frac{\partial^2 H}{\partial q_r \partial p} \delta q_r + \sum \frac{\partial^2 H}{\partial p_r \partial p} \delta p_r. \quad (40)$$

This follows straight from the equation, $\dot{q} = \partial H / \partial p$. If we substitute for δq_r and δp_r by means of

$$\delta q_r = \varepsilon \frac{\partial \phi}{\partial p_r}, \quad \delta p_r = -\varepsilon \frac{\partial \phi}{\partial q_r} \quad (41)$$

where ε is a small constant,

$$\begin{aligned} \frac{d}{dt} \delta q &= \varepsilon \left\{ \sum \frac{\partial^2 H}{\partial q_r \partial p} \cdot \frac{\partial \phi}{\partial p_r} - \sum \frac{\partial^2 H}{\partial p_r \partial p} \frac{\partial \phi}{\partial q_r} \right\} \\ &= \varepsilon \frac{\partial}{\partial p} \left\{ \sum \frac{\partial H}{\partial q_r} \frac{\partial \phi}{\partial p_r} - \sum \frac{\partial H}{\partial p_r} \frac{\partial \phi}{\partial q_r} \right\} - \sum \frac{\partial H}{\partial q_r} \frac{\partial^2 \phi}{\partial p_r \partial p} + \sum \frac{\partial H}{\partial p_r} \frac{\partial^2 \phi}{\partial q_r \partial p} \\ &= \varepsilon \frac{\partial}{\partial p} \left\{ -\sum \dot{p}_r \frac{\partial \phi}{\partial p_r} - \sum \dot{q}_r \frac{\partial \phi}{\partial q_r} \right\} + \sum \dot{p}_r \frac{\partial^2 \phi}{\partial p_r \partial p} + \sum \dot{q}_r \frac{\partial^2 \phi}{\partial q_r \partial p} \\ &= \varepsilon \left[\frac{\partial}{\partial p} \left(\frac{\partial \phi}{\partial t} - \frac{d\phi}{dt} \right) + \frac{d}{dt} \left(\frac{\partial \phi}{\partial p} \right) - \frac{\partial}{\partial t} \left(\frac{\partial \phi}{\partial p} \right) \right] \\ &= \varepsilon \frac{d}{dt} \left(\frac{\partial \phi}{\partial p} \right), \end{aligned}$$

since $d\phi/dt$ vanishes on account of ϕ being integral. Thus variational equations for δq and δp are satisfied by (41). Now the transformation from q_r, p_r to \dot{q}_r, \dot{p}_r where

$$\dot{q}_r = \dot{q}_r + \delta q_r, \quad \dot{p}_r = \dot{p}_r + \delta p_r \quad (43)$$

is a contact transformation when a function $W(q, p)$ exists such that

$$\sum (\dot{p}_r d\dot{q}_r - \dot{p}_r dq_r) = dW \quad (44)$$

$$\begin{aligned} \text{or} \quad & \Sigma \{ (p_r + \delta p_r) (dq_r + \delta dq_r) - p_r dq_r \} = dW \\ \text{or} \quad & \Sigma (dq_r \delta p_r + p_r \delta dq_r) = dW, \end{aligned} \quad (45)$$

neglecting the term $\delta p_r \delta dq_r$ of the higher order. Let

$$\delta p_r = \epsilon \dot{p}_r, \quad \delta dq_r = \epsilon \dot{q}_r. \quad (46)$$

If $W = u\epsilon$ and ϵ is a small constant,

$$\Sigma (dq_r \dot{p}_r + p_r \dot{q}_r) = du \quad (47)$$

$$\text{or} \quad \Sigma (\dot{p}_r dq_r - \dot{p}_r \dot{q}_r) = d(u - \Sigma p_r \dot{q}_r) \equiv dk. \quad (48)$$

$$\text{Hence} \quad \dot{p} = \frac{\partial k}{\partial q}, \quad \dot{q} = -\frac{\partial k}{\partial p}. \quad (49)$$

By comparing (41) with (49) we find that (41) is a contact transformation. All infinitesimal contact transformations which preserve the physical content of the system, that is, which transform the family of orbits into itself are of the type (41). On the other hand, the transformation which replaces H_0 by $H_0 + R = \dot{H}$ even where R is an infinitesimal is not in general of the type (41). Hence the dynamical system characterised by H_0 is different from that characterised by \dot{H} . It is the possibility of going from one dynamical system to another, through a contact transformation, which enhances the value of the latter.

As a dynamical system progresses the values of the variables q_r, p_r at time t are replaced by

$$q_r + \frac{\partial H}{\partial p_r} \delta t, \quad p_r - \frac{\partial H}{\partial q_r} \delta t, \quad (50)$$

at $t + \delta t$. H being an integral of the system the motion may be interpreted as a progressive transition from one set of variables to another through infinitesimal contact transformations. The same fact is expressed alternatively when we say that Hamilton's function $\int_{t_1}^t L dt$ is a solution of the Hamilton-Jacobi differential equation. *

4. POISSON BRACKETS

The Poisson brackets¹² play a very conspicuous part in classical as well as quantum mechanics. The non-commutative relations of the latter theory were deduced by Dirac from the Poisson brackets expressed in action-angle variables. The same relations, which had appeared first in a paper by Heisenberg, were suggested by the matrix-formulation of observables. As the classical Poisson brackets of a dynamical system satisfy conditions which are very similar in form to the quantum conditions some important properties of these classical expressions will be given here. *

Consider a dynamical system characterised by the n co-ordinates q_r and the n momenta p_r , r running from 1 to n . If u, v are any two functions of the co-ordinates and the momenta, the Poisson bracket $[u, v]$ is given by

$$[u, v] = \sum_{r=1}^n \left(\frac{\partial u}{\partial q_r} \frac{\partial v}{\partial p_r} - \frac{\partial u}{\partial p_r} \frac{\partial v}{\partial q_r} \right). \quad (51)$$

Hence $[u, v] = -[v, u]$ and $[u, u] = 0$

and it can be easily verified that if w is another function of q 's and p 's

$$[(u, v), w] + [(v, w), u] + [(w, u), v] = 0. \quad (52)$$

This is a well known identity due to Jacobi. Directly from the definition of the Poisson bracket the following obvious relations follow :

$$\begin{aligned} [q_r, p_r] &= 1, & [q_r, p_s] &= 0, \\ [p_r, p_s] &= 0, & [q_r, q_s] &= 0, \end{aligned} \quad \begin{aligned} r, s &= 1, 2, 3, \dots, n \\ r &\neq s. \end{aligned} \quad (53)$$

The Hamiltonian equations of motion are equivalent to

$$\dot{q} = [q, H], \quad \dot{p} = [p, H] \quad (54)$$

where, as usual, H denotes the Hamiltonian. Further, if the variables are changed by a contact transformation so that the equations of motion now become

$$\dot{q}' = [q', H'], \quad \dot{p}' = [p', H'] \quad (55)$$

the two sets of variables are connected by the relations,

$$\begin{aligned} [\dot{q}_r, \dot{p}_r] &= 1, & [\dot{q}_r, \dot{p}_s] &= 0, \\ [\dot{p}_r, \dot{p}_s] &= 0, & [\dot{q}_r, \dot{q}_s] &= 0, \end{aligned} \quad \begin{aligned} r, s &= 1, 2, \dots, n \\ r &\neq s. \end{aligned} \quad (56)$$

It may be emphasized that (53) are obvious identities while (56) are equations connecting q - p and \dot{q} - \dot{p} . For establishing (56) it is enough to show that for an infinitesimal contact transformation, viz.,

$$\dot{q}_r = q_r + \epsilon \frac{\partial \phi}{\partial p_r}, \quad \dot{p}_r = p_r - \epsilon \frac{\partial \phi}{\partial q_r}, \quad (57)$$

where ϕ is any function of p - q , the relations are valid if ϵ^2 is negligible. Inasmuch as every contact transformation may be supposed to be generated by successive infinitesimal contact transformations the relations (56) become valid generally. The usual proof of (56) is based on the notion of Lagrange brackets.

An immediate result of (56) is the important property of Poisson brackets usually stated explicitly as under :

$$[u, v] = \sum \left(\frac{\partial u}{\partial q_r} \frac{\partial v}{\partial p_r} - \frac{\partial u}{\partial p_r} \frac{\partial v}{\partial q_r} \right) = \sum \left(\frac{\partial u}{\partial \dot{q}_r} \frac{\partial v}{\partial \dot{p}_r} - \frac{\partial u}{\partial \dot{p}_r} \frac{\partial v}{\partial \dot{q}_r} \right). \quad (58)$$

which can be directly verified by using the values of the n^2 brackets (56). In general

$$\begin{aligned}\frac{du}{dt} &= \frac{\partial u}{\partial t} + \sum \frac{\partial u}{\partial q_r} [q_r, H] + \sum \frac{\partial u}{\partial p_r} [p_r, H] \\ &= \partial u / \partial t + [u, H].\end{aligned}\quad (59)$$

Hence if t does not appear explicitly in u the condition for u to be an integral of the system is

$$[u, H] = 0. \quad (60)$$

By using the identity (52), $[u, v]$ may be shown to be an integral of the system if u and v are any two independent integrals.

Let us now expand a Poisson bracket of the form $[xy, zw]$ without disturbing the order of x relative to y and of z relative to w . Two expansions¹⁴ become possible.

$$\begin{aligned}[xy, zw] &= x[y, zw] + [x, zw]y \\ &= x\{[y, z]w + z[y, w]\} + \{[x, z]w + z[x, w]\}y.\end{aligned}\quad (61)$$

Similarly

$$\begin{aligned}[xy, zw] &= [xy, z]w + z[xy, w] \\ &= \{x[y, z] + [x, z]y\}w + z\{x[y, w] + [x, w]y\}.\end{aligned}\quad (62)$$

Comparing the two expansions we get

$$xz[y, w] - zx[y, w] = [x, z]yw - [x, z]wy,$$

or,

$$(xz - zx)[y, w] = [x, z](yw - wy). \quad (63)$$

In classical mechanics $xz=zx$, $yw=wy$ but in quantum mechanics where the variables have a different significance this is not always true. By a postulate of Dirac's theory,

$$\begin{aligned}(xz - zx) &= \frac{i\hbar}{2\pi} [x, z], \\ (yw - wy) &= \frac{i\hbar}{2\pi} [y, w],\end{aligned}\quad (64)$$

\hbar being Planck's constant and i , a square root of -1. The trivial relations (53) now become the quantum conditions of Dirac's theory :

$$\begin{aligned}q_r \hat{p}_r - \hat{p}_r q_r &= \frac{i\hbar}{2\pi}, \quad q_r \hat{p}_s - \hat{p}_s q_r = 0, \\ \hat{p}_r \hat{p}_s - \hat{p}_s \hat{p}_r &= 0, \quad q_r q_s - q_s q_r = 0, \\ r, s &= 1, 2, \dots, n \\ r &\neq s,\end{aligned}\quad (65)$$

It is worthwhile noting that (63) becomes an identity also when the values of Poisson brackets as given by (64) are substituted therein. These (65) are also known as exchange relations.

5. CONDITIONALLY PERIODIC SYSTEMS

Consider a conservative dynamical system for which it is possible to use co-ordinates splitting the Hamilton-Jacobi equation given by the energy integral into a series of equations of the form¹⁶

$$\left(\frac{\partial F}{\partial q_i}\right)^2 = f_i(q_i, \alpha_1, \alpha_2, \dots, \alpha_n). \quad (66)$$

We have in such a case

$$F = \sum_{i=1}^n \int (f_i)^{1/2} dq_i. \quad (67)$$

Further let us suppose that when f_i does not reduce to a constant there exist two excessive roots \dot{q}_i and \ddot{q}_i between which q_i must lie if $\partial F/\partial q_i$ is to be real. In the latter case there is a libration in the q_i co-ordinate between \dot{q}_i and \ddot{q}_i . We define for a complete libration in the q_i co-ordinate the integral I_i as given by

$$I_i = \oint p_i dq_i = \oint (f_i)^{1/2} dq_i. \quad (68)$$

During a complete libration of q_i , F increases by I_i which is therefore called a modulus of periodicity of F . It is clear that each I is a function of α 's since \dot{q} and \ddot{q} , the limits of integration, are roots of

$$f(q, \alpha_1, \alpha_2, \dots, \alpha_n) = 0.$$

Let us suppose that there are as many independent I 's as there are α 's. F can now be expressed as a function of q 's and I 's. If we denote $\partial F/\partial I$ by W there is the contact transformation

$$p = \frac{\partial F}{\partial q}, \quad W = \frac{\partial F}{\partial I}, \quad H' = H, \quad (69)$$

$\partial F/\partial t$ being zero. The new Hamiltonian equations are

$$\frac{\partial H}{\partial I} = \dot{W}, \quad \frac{\partial H}{\partial W} = -\dot{I}. \quad (70)$$

When the energy integral $H_0 = \alpha_1$ exists H_0 is expressible as a function of I 's only. Therefore,

$$\frac{\partial H}{\partial W} = 0 \quad (71)$$

and I is a constant. Moreover $\partial H/\partial I$, being a function of I 's only, is constant for the motion. So

$$\dot{W} = \omega, \quad W = \omega t + s, \quad (72)$$

the constants ω and s being defined by the equations themselves. The variables I - W are the action-angle variables of the classical theory and ω 's are frequencies.

The relations between I - W and q are best brought out by showing the effect of a libration in q_i on W_k . For

$$\begin{aligned} \Delta W_k &\equiv \oint \frac{\partial W_k}{\partial q_i} dq_i = \oint \frac{\partial^2 F}{\partial I_k \partial q_i} dq_i \\ &= \frac{\partial}{\partial I_k} \oint \frac{\partial F}{\partial q_i} dq_i = \frac{\partial}{\partial I_k} I_i = \delta_k^i; \end{aligned} \quad (73)$$

that is, the effect on W_k is zero or an increase by unity according as the coordinate of libration is other than q_k or q_k itself. Hence the q 's are functions of W , I being periodic in each W with period unity. Thus

$$q = \sum \dots C \tau_1 \tau_2 \dots \tau_n e^{2\pi i (\tau_1 \omega_1 + \tau_2 \omega_2 + \dots + \tau_n \omega_n) t} \quad (74)$$

where $\tau_1, \tau_2, \dots, \tau_n$ are integers or zero and the coefficients $C \tau_1 \dots \tau_n$ are functions of I_1, I_2, \dots, I_n . If there exists a set of integers for which

$$\tau_1 \omega_1 + \tau_2 \omega_2 + \dots + \tau_n \omega_n = 0 \quad (75)$$

the system is said to be degenerate. It may be singly or multiply degenerate according to the number of distinct relations like (75) which connect the frequencies. A non-degenerate system evidently will not pass through the same configuration twice but it can be brought as near as one likes to any configuration of the past by choosing a sufficiently distant epoch. Hence a non-degenerate system is called conditionally periodic. For a doubly periodic system, for example, let

$$q = \sum \sum C \tau_1 \tau_2 e^{2\pi i (\tau_1 \omega_1 + \tau_2 \omega_2) t}$$

where $\omega_1 = \pi \omega_2$. When $t = 0$,

$$q = q_0 = \sum \sum C \tau_1 \tau_2.$$

For no other value of t , $q = q_0$. But if we take $\omega_1 = 3 \omega_2$, approximately, q is nearly equal to q_0 when $t = 1/\omega_2$. Similarly if we take $\omega_1 = 31/\omega_2$ as a better approximation, we have to consider the epoch $t = 10/\omega_2 = 31/\omega_1$ for getting q nearly equal to q_0 .

A relation of great interest in Bohr's theory may now be obtained. The Hamiltonian of a conservative system is equal α_1 and hence a function of I 's only as already stated above. Hence by (70) and (72), if the variation in energy is δE ,

$$\delta E = \sum \omega \delta I. \quad (76)$$

By the conditions of the contact transformation (69)

$$\Sigma p \delta q + \Sigma W \delta I = \delta F$$

or,

$$\Sigma p \delta q - \Sigma I \delta W = \delta (F - \Sigma IW) = \delta F'. \quad (77)$$

For a complete libration in any of the q 's there is an increase of the corresponding I in F as also in the function ΣIW as (68) and (73) show. Hence F' remains unchanged in a complete libration. This fact is used to prove that the mean kinetic energy over a long interval of time is $\frac{1}{2} \Sigma I \omega$. For the required mean over a long period of time t_0 is, by (70)

$$\begin{aligned} \frac{1}{2} \int_0^{t_0} \Sigma p \dot{q} dt / t_0 &= \frac{1}{2 t_0} \int_0^{t_0} (\Sigma I \dot{W} + \dot{F}') dt \\ &= \frac{1}{2} \Sigma I \omega + \frac{1}{2 t_0} [F']_0^{t_0}. \end{aligned} \quad (78)$$

$[F']_0^{t_0}$ is bound to remain finite because of what is known of F' . As t_0 increases the average value of the kinetic energy over a long period is found to be very nearly $\frac{1}{2} \Sigma I \omega$. It is only in the limit as $t_0 \rightarrow \infty$ that this value is actually attained.

The action functions I play a very important part in Bohr's theory. This is due to the following important theorem leading to Ehrenfest's¹⁶ adiabatic principle :

"The I -values are the same for two dynamical systems which are distinguished only by different values of a parameter in the Hamiltonians."

The theorem is not true for degenerate systems.

To prove the theorem consider the Hamiltonian $H(q, p, c)$ in which c is the parameter varying with time. The contact transformation (77) gives

$$\frac{\partial F'}{\partial q} = p, \quad \frac{\partial F'}{\partial W} = -I, \quad \dot{H}' = H + \frac{\partial F'}{\partial t} \quad (79)$$

F' containing t explicitly in this case. There follows the equation

$$\frac{dI}{dt} = -\frac{\partial \dot{H}'}{\partial W} \equiv -\frac{\partial H}{\partial W} - \frac{\partial^2 F'}{\partial W \partial t}. \quad (80)$$

W does not appear in H in the case when c is a constant. The effect of a varying c is only to make the I 's functions of t . Therefore,

$$\frac{\partial H}{\partial W} = 0 \quad (81)$$

and
$$\frac{\partial F'}{\partial t} = \frac{\partial F'}{\partial c} \dot{c} = \dot{c} \Sigma A_\tau e^{2\pi i (\tau_1 \omega_1 + \dots + \tau_n \omega_n) t}. \quad (82)$$

Hence
$$\frac{\partial^2 F'}{\partial W \partial t} = \dot{c} \Sigma B_\tau e^{2\pi i (\tau_1 \omega_1 + \dots + \tau_n \omega_n) t} \quad (83)$$

there being no constant term, as a result of differentiation, in the right-hand side of (83). In writing the expansion of F' we have used the earlier result that F' is a multiply periodic function when c is a constant. We expect A_{τ} , B_{τ} to be coefficients involving t through c . ω 's also will involve t through c . In spite of these complications it is obvious from (80), (81) and (83) that

$$\left[I \right]_{t_1}^{t_2} = - \int_{t_1}^{t_2} \frac{\partial^2 F'}{\partial \omega \partial t} dt = - \dot{c}_m \alpha \quad (84)$$

where we take \dot{c}_m as the mean value α being a finite multiplying factor. Hence by making \dot{c}_m conceptually small enough we get

$$\left[I \right]_{t_1}^{t_2} = 0. \quad (85)$$

The I -values are therefore invariant for a change in the value of the parameter c .

By way of illustration an oscillator may be considered, given by

$$\ddot{x} + \mu x = 0.$$

When $\mu = \mu_0$ let $x = x_0 \cos (p_0 t + \epsilon_0)$

where $\mu_0 = p_0^2$.

When $\mu = \mu_1$ let $x = x_1 \cos (p_1 t + \epsilon_1)$

where $\mu_1 = p_1^2$.

Let us suppose that the transition from $\mu = \mu_0$ to $\mu = \mu_1$ takes place very slowly and that during the transition

$$\mu = \mu(t), \quad p = p(t), \quad a = a(t), \quad \epsilon = \epsilon(t).$$

Substituting $x = a \cos (pt + \epsilon)$

in the differential equation we find

$$\left[\ddot{a} - a \left\{ \frac{d}{dt} (pt + \epsilon) \right\}^2 \right] = -\mu a,$$

or

$$a^2 \frac{d}{dt} (pt + \epsilon) = C.$$

Hence if a , p , ϵ are very slowly varying functions of time we get what is analogous to (85)

$$a^2 p = C, \text{ or } a_0^2 p_0 = a_1^2 p_1 = C. \quad (85')$$

It is this result, as will be shown later, which makes Planck's hypothesis so successful and earns for it a central position in the logical development of the quantum theory.

6. DEGENERACY

Let the frequencies of a multiply periodic system¹⁵ be connected by the relation of degeneracy,

$$\tau_1 \omega_1 + \tau_2 \omega_2 + \dots + \tau_n \omega_n = 0. \quad (86)$$

This is equivalent to

$$\tau_1 \frac{\partial H}{\partial I_1} + \tau_2 \frac{\partial H}{\partial I_2} + \dots + \tau_n \frac{\partial H}{\partial I_n} = 0. \quad (87)$$

The general integral of the partial differential equation (87) is

$$H = \phi(J_1, J_2, \dots, J_{n-1}) \quad (88)$$

where $J_1 = \frac{I_1}{\tau_1} - \frac{I_n}{\tau_n}, J_2 = \frac{I_2}{\tau_2} - \frac{I_n}{\tau_n}, \text{ etc. etc.}$

It will be remembered that $\tau_1, \tau_2, \text{ etc.}$, are all integers. The J 's are the new independent action variables. Their number is one less than that of the I 's. When the number of independent action variable is less than that of coordinates the system is degenerate. Each condition of degeneracy reduces the number of independent action functions by one. The angle variables are given by

$$v = \frac{\partial F}{\partial J}, \quad (90)$$

where F is the function of the contact transformation. The connections between W 's and v 's are given by

$$\begin{aligned} W_1 &= \frac{\partial F}{\partial I_1} = \frac{\partial F}{\partial J_1} \frac{\partial J_1}{\partial I_1} + \frac{\partial F}{\partial J_2} \frac{\partial J_2}{\partial I_1} + \dots + \frac{\partial F}{\partial J_{n-1}} \frac{\partial J_{n-1}}{\partial I_1}, \\ &= v_1/\tau_1. \end{aligned} \quad (91)$$

Similarly,

$$\begin{aligned} W_2 &= v_2/\tau_2, \text{ etc.}; \quad W_{n-1} = v_{n-1}/\tau_{n-1}, \\ W_n &= \frac{\partial F}{\partial I_n} = - (v_1 + v_2 + \dots + v_{n-1})/\tau_n. \end{aligned} \quad (92)$$

The coordinates and the momenta of the system can be expressed in terms of J 's and v 's by using the above results.

Let us work out fully a simple degenerate case.

For an electron moving in the field of a proton,

$$T = \frac{1}{2} m_e (\dot{r}^2 + r^2 \dot{\theta}^2), \quad V = -e^2/r$$

in the usual notation. As

$$p_1 = \frac{\partial T}{\partial \dot{r}}, \quad p_2 = \frac{\partial T}{\partial \dot{\theta}}.$$

the Hamiltonian is

$$H = \frac{1}{2m} \left(p_1^2 + \frac{p_2^2}{r^2} \right) - \frac{e^2}{r}.$$

The Hamilton-Jacobi equation is

$$\frac{1}{2m} \left[\left(\frac{\partial F}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial F}{\partial \theta} \right)^2 \right] - \frac{e^2}{r} = \alpha_1.$$

Hence

$$F = \alpha_2 \theta + \int \left(2m \alpha_1 + \frac{2me^2}{r} - \frac{\alpha_2^2}{r^2} \right)^{\frac{1}{2}} dr.$$

$$\begin{aligned} I_1 &= \oint p_1 dr = \oint \frac{\partial F}{\partial r} dr \\ &= \oint \frac{dr}{r} \left\{ A (r_1 - r) (r - r_2) \right\}^{\frac{1}{2}} = 2\pi \left(\frac{me^2}{(-2m\alpha_1)^{\frac{1}{2}}} - \alpha_2 \right), \end{aligned}$$

where

$$r_1 + r_2 = -e^2/\alpha_1, \quad r_1 r_2 = -\alpha_2^2/2m\alpha_1, \quad A = -2m\alpha_1,$$

and

$$\begin{aligned} I_2 &= \oint p_2 d\theta = 2\pi \alpha_2, \\ \alpha_1 &= H = -\frac{2\pi^2 m e^4}{(I_1 + I_2)^2}. \end{aligned}$$

The total energy is expressed as a function of I_1 and I_2 only. The frequencies are given by

$$\omega_1 = \frac{\partial H}{\partial I_1}, \quad \omega_2 = \frac{\partial H}{\partial I_2}.$$

As $\omega_1 - \omega_2 = 0$ the dynamical system is degenerate. Obviously we can write

$$H = -\frac{2\pi^2 m e^4}{J_1^2},$$

if we take

$$J_1 = I_1 + I_2.$$

The angle variable corresponding to J_1 is given by

$$W_1 = \frac{\partial F}{\partial I_1} = \frac{\partial F}{\partial J_1} \cdot \frac{\partial J_1}{\partial I_1} = v_1.$$

7. RELATIVISTIC EXPRESSIONS

The vector¹⁷ associated with the motion of a particle in special relativity is

$$mc \frac{dx}{ds}, \quad mc \frac{dy}{ds}, \quad mc \frac{dz}{ds}, \quad mc^2 i \frac{dt}{ds} \quad (93)$$

where m is the rest mass of the particle, c the velocity of light, $i^2 = -1$ and

$$ds^2 = c^2 dt^2 - dx^2 - dy^2 - dz^2. \quad (94)$$

Hence the components of the vector are, if $v^2 = \dot{x}^2 + \dot{y}^2 + \dot{z}^2$,

$$\frac{m\dot{x}}{(1-v^2/c^2)^{\frac{1}{2}}}, \quad \frac{m\dot{y}}{(1-v^2/c^2)^{\frac{1}{2}}}, \quad \frac{m\dot{z}}{(1-v^2/c^2)^{\frac{1}{2}}}, \quad \frac{mc i}{(1-v^2/c^2)^{\frac{1}{2}}}, \quad (95)$$

or $p_1, p_2, p_3, iE/c$,

where p_1, p_2, p_3 is the momentum vector and E is the Energy.

$$p_1^2 + p_2^2 + p_3^2 - E^2/c^2 = m^2 c^2, \quad (96)$$

$$E = mc^2 / (1 - v^2/c^2)^{\frac{1}{2}}. \quad (97)$$

For a photon or a quantum of radiation $v = c$ and its energy can be finite only if the rest mass m vanishes. The condition satisfied by a photon is

$$E^2 = c^2 (p_1^2 + p_2^2 + p_3^2). \quad (98)$$

The Lagrangian $L \equiv T - V$ must be such that, for a particle,

$$\frac{\partial L}{\partial \dot{x}} = p_1, \quad \frac{\partial L}{\partial \dot{y}} = p_2, \quad \frac{\partial L}{\partial \dot{z}} = p_3. \quad (99)$$

This is achieved by taking

$$L = mc^2 \left[1 - \left(1 - v^2/c^2 \right)^{\frac{1}{2}} \right] - V. \quad (100)$$

It is worth pointing out that the expression which replaces the classical kinetic energy in L is different from the relativistic expression of the kinetic energy. This latter is given by

$$T = mc^2 \left[\left(1 - v^2/c^2 \right)^{-\frac{1}{2}} - 1 \right]. \quad (101)$$

For

$$\begin{aligned} \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) &= - \frac{\partial V}{\partial x}, \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{y}} \right) &= - \frac{\partial V}{\partial y}, \\ \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{z}} \right) &= - \frac{\partial V}{\partial z}. \end{aligned}$$

This gives

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}} \right) \cdot \dot{x} + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{y}} \right) \cdot \dot{y} + \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{z}} \right) \cdot \dot{z} + \frac{dV}{dt} = 0.$$

The resulting integral is

$$\frac{mc^2}{(1 - v^2/c^2)^{\frac{1}{2}}} + V = \text{const.} \quad (102)$$

The kinetic energy is defined as in (101) firstly because it agrees with the expression of energy as in (102) and secondly for the reason that as $c \rightarrow \infty$ the classical expression of kinetic energy is realised from (101).

8. THE FAILURE OF CLASSICAL MECHANICS

The successes of the theory need not be recorded here. They are found in practically every branch of physical science although the most striking and sensational of the successes are claimed by astronomers who have made accurate predictions of comets and planets. Discussing the failure of the theory we will not be concerned with astronomical situations such as the motion of the perihelion of Mercury as that will entail a digression into general relativity. About the beginning of the present century three major problems of atomic physics had defeated all attempts of solution on classical lines. They may be briefly mentioned here. Firstly, the atoms and molecules must be very stable structures if substances are to possess definite physical and chemical properties; further, their internal structure must consist of charges in motion as otherwise their characteristic radiations cannot be accounted for. No model of the atom could reconcile the stability, which was taken for granted in the kinetic theory of gases, with the motion of the inner charges as given by classical electrodynamics and mechanics. Secondly, it was a mere matter of faith that the statistical laws which a complicated system such as a gas obeys could be deduced from classical laws. This deduction was never accomplished and the view was gaining ground that the laws of large numbers in physics remain outside the ambit of classical mechanics. The third and the most discussed problem was that of heat radiation. An account of it is given in a discussion of the quantum hypothesis which removed the difficulties of this problem and explained a host of phenomena linked up with it such as that of unaccountably low specific heats at low temperatures.

REFERENCES

1. THOMSON and TAIT, *Natural Philosophy*, Vol. I, Part 1, pp. 240-9 (1879).
2. MACH, E., *Science of Mechanics*, p. 340, (1893).
3. THOMSON and TAIT, *Loc. Cit.*, Vol. I, Part 2, p. 9 (1883).
4. WEYL, H., *Mind and Nature*, pp. 60-80 (1934).
- 5a. OSGOOD, W., *Mechanics*, pp. 345-55 (1937).
- 5b. PEARSON, K., *The Grammar of Science*, pp. 235-75 (1937).
6. HARTREE, D. R., *Mathematical Gazette*, Vol. 19, pp. 73-8 (1935).
7. MISES, R. VON., *Probability, Statistics and Truth*, pp. 304-7 (1939).
8. RUSSELL, B., *The Principles of Mathematics*, p. 348 and pp. 482-3 (1937).
9. SCHRÖDINGER, E., *Science and the Human Temperament*, pp. 43-65 (1935).

10. CONDON and SHORTLY, *The Theory of Atomic Spectra*, p. 429 (1935).
 - 11a. OSGOOD, W., *Loc. Cit.*, pp. 356-88 (1937).
 - 11b. MAXWELL, J. C., *Matter and Motion*, pp. 145-61 (1920).
 - 12a. WHITTAKER, E. T., *Analytical Dynamics*, pp. 288-355 (1927).
 - 12b. TOLMAN, R. C., *Principles of Statistical Mechanics*, pp. 16-42 (1938).
 13. EISENHART, L. P., *Continuous Groups of Transformations*, pp. 239 (1933).
 14. DIRAC, P. A. M., *Quantum Mechanics*, pp. 88-91 (1935).
 15. BIRTWISTLE, G., *The Quantum Theory of the Atom*, pp. 56-79 (1926).
 - 16a. JEANS, J. H., *The Dynamical Theory of Gases*, pp. 406-15 (1921).
 - 16b. RUARK and UREY, *Atoms Molecules and Quanta*, pp. 176-9 (1930).
 - 17a. BIRTWISTLE, G., *Loc. Cit.*, pp. 48-52 (1926).
 - 17b. FRENKEL, J., *Wave Mechanics*, Part 1, pp. 9-14 (1936).
-

II THE QUANTUM HYPOTHESIS

1. The Origin of the Quantum Hypothesis. 2. Einstein's Relations.
3. The Success of the Quantum Hypothesis.

1. THE ORIGIN OF THE QUANTUM HYPOTHESIS

Atoms have been talked of since the days of Democritus and Lucretius but the modern theory of discontinuous matter owes its origin to Dalton. Planck and Einstein¹ are jointly responsible for the quantum hypothesis which asserts the discontinuity of energy. The work of a series of investigators starting from Faraday's laws of electrolysis has culminated in the postulate of the electron which implies the discontinuity of electric charge. On a survey of the historical development the quantum hypothesis does not appear to-day as anything bizarre or revolutionary; on the contrary it appears as a logical sequel to the earlier hypothesis of Dalton and in keeping with the speculations that were current in 1900 about the atomocity of charge.

The physical situation which first suggested a mathematical formulation of the quantum postulate originated in the topic of the temperature radiation of a black body.

A black body is an ideal body which absorbs all incident radiation there being no reflection, scattering or transmission. Lamp black approaches this ideal; so also a hollow metal globe with a small hole in it as the chance of escape for any radiation entering it is very small. In general, a body which is kept at a given temperature may emit a characteristic radiation without undergoing a structural change. This characteristic radiation is a pure temperature radiation since it depends upon the temperature alone. Theoretically the study of temperature radiations is simplest for a black body. Classical mechanics and the law of equipartition of energy given by classical statistics led to Rayleigh's formula, viz.,

$$I(\lambda, T) = c T \lambda^{-4}, \quad (1)$$

although Wien had derived from the classical thermodynamical principles and a speculative assumption of a Maxwellian distribution of velocities the alternative formula, viz.,

$$I(\lambda, T) = a \lambda^{-5} e^{-b/\lambda T}, \quad (2)$$

here $I(\lambda, T)$ is the intensity of radiation for wave lengths between λ and $\lambda + d\lambda$ at temperature T and a, b, c are certain constants. Rayleigh's

formula was found to be true for large wave lengths but it gives very large values of $I(\lambda, T)$ for small values of λ producing an appreciable divergence between theory and experiment. The other formula agrees beautifully with experimental results only for small values of λT but for large values of the same the intensity function becomes too large. A generalized formula was required which would approach (2) for small values of λT , (1) for large values of λT and which would not violate the well-established Stephen's law of radiation. By means of his hypothesis Planck succeeded in deriving the successful formula :

$$I(\lambda, T) = a \lambda^{-5} / (e^{b/\lambda T} - 1). \quad (3)$$

It is in full accord with experimental results. It escapes the ultra-violet catastrophe produced by the Rayleigh formula and the infra-red disaster due to the other.

In his derivation of (3) Planck assumed that the mechanism responsible for radiation consists of linear oscillators performing simple harmonic motion. A particle of mass m executing simple harmonic motion in a straight line satisfies the equation,

$$m\ddot{x} + \mu x = 0, \quad (4)$$

the force per unit mass being attractive and proportional to the distance from the origin. So

$$x = a \cos \sqrt{\mu/m} t \quad (5)$$

if the time is reckoned from a suitable epoch. The total energy at any instant is given by

$$E = \frac{1}{2} m \dot{x}^2 + \frac{1}{2} \mu x^2, \quad (6)$$

which is constant by the principle of energy. Denoting the frequency of the periodic motion by ν and hence the period by $1/\nu$ we have

$$\frac{1}{\nu} = \frac{2\pi}{\sqrt{\mu/m}}. \quad (7)$$

Hence from (6) and (7)

$$E = A\nu \quad (8)$$

where

$$A = 2\pi^2 m a^2 \nu. \quad (9)$$

It is interesting to note that as

$$p = m \dot{x} \\ \oint p dx = \int_0^{2\pi/\sqrt{(\mu/m)}} \mu a^2 \sin^2 \sqrt{\mu/m} t dt = A. \quad (10)$$

In the problem of heat radiation the linear oscillators, representing the effect of a black body in equilibrium with radiation at a specific temperature, were supposed to exist in discrete states of energy so that

$$A = nh, \quad (11)$$

where h is Planck's constant and n is a non-negative integer. Thus A could not be a continuous function and, in a state of equilibrium, energy was assumed to be absorbed or emitted in discrete quanta of multiples of $h\nu$. h , of course, has the small value 6.5×10^{-27} in c. g. s. units. The assumption (11) gave (3).

Originally, the assumption was made for no other reason except that of mathematical convenience and enough justification was found for it in the result (3) and in several others. What is curious, however, is that the expression for energy should have been split up in the particular manner of (8) and (9). The explanation lies in the significance of A . The area of the ellipse obtained by plotting the position and the momentum in one period in perpendicular directions is A , as the result (10) shows; and the far reaching consequences of Planck's hypothesis are due to the fact that A is one of the action invariants discussed in (I, 85-85'). As all conservative systems of classical mechanics are governed by a continuous action function, the implication of Planck's hypothesis that atomic phenomena are similarly governed by the quantum of action h , appearing in (11), augured very well for subsequent research.²

The integer n is an important feature of Planck's hypothesis. Kronecker once said that God made integers the rest being the work of man. We do not know to what extent Planck was under the influence of Kronecker's dictum when he thought of the famous hypothesis. A pure mathematician intimately acquainted with prime numbers, might have insisted that n must be a prime number. There is no *a priori* reason why n should be an integer. It cannot however be denied that, with our habits of thought and with our experience of nature, a 'simpler' statement of the occurrence of discrete values of energy is not possible.

The full content of the quantum hypothesis is not described by (11). The existence of discrete states of energy for oscillators merely implies that absorption and emission are not continuous as far as their effects are observable. The more daring hypothesis was suggested by Einstein in 1905 that radiation itself is discontinuous; the energy E of a radiation with frequency ν was supposed to be carried by unidirectional quanta satisfying the relation,

$$h\nu = E. \quad (12)$$

Ten years of experimental work on the photoelectric effect of light had to pass before (12) was accepted. In 1928 Lawrence and Beams tried to

measure the lengths of the unidirectional quanta of radiation. They concluded that the lengths must be less than 10^{-9} cm.

Einstein's hypothesis (12) is vitally connected with another fundamental relation usually attributed to him, viz.,

$$E = mc^2 \quad (12)$$

where E is the energy and m , the rest mass of a particle. The atomicity of matter is naturally expected therefore to imply the atomicity of energy.

An immediate application of Planck's formula of heat radiation was made by Einstein to modify, as experimental results required, the Dulong and Petit law of atomic heats. The precise form of Planck's law (3) is

$$I(\lambda, T) = \frac{hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{k\lambda T}} - 1}, \quad (13)$$

in the usual notation. Since

$$\lambda\nu = c \quad (14)$$

the density of radiation, $J(\nu, T)$, in the frequency range is given by

$$\begin{aligned} J(\nu, T) d\nu &= -I(\lambda, T) d\lambda \\ &= \frac{h\nu^5}{c^3} \cdot \frac{c}{\nu^2} d\nu \frac{1}{e^{\frac{h\nu}{kT}} - 1} \end{aligned}$$

$$\text{so that} \quad J(\nu, T) = \frac{h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kT}} - 1}. \quad (15)$$

The mean energy of an oscillator \bar{W} is connected with this expression by

$$\bar{W} = \frac{c^2}{\nu^3} J(\nu, T) = \frac{h\nu}{e^{\frac{h\nu}{kT}} - 1}. \quad (16)$$

For oscillators with three degrees of freedom the mean energy would be thrice this amount. Hence the heat content per gramme-mole is

$$\bar{E} = \frac{3Nh\nu}{e^{\frac{h\nu}{kT}} - 1}, \quad (17)$$

N being Avogadro's number. For the atomic heat at constant volume we get the formula, first given by Einstein,

$$C_v = \frac{d\bar{E}}{dT}. \quad (18)$$

According to the Dulong and Petit law the atomic heats of all substances are the same at all temperatures. It had been discovered by experimenters that the atomic heats are usually much smaller at low temperatures.

and this fact could be explained by using Planck's formula for the mean energy of an atom in place of the law of equipartition of energy. By virtue of the latter

$$E = 3 Nk T$$

which gives a constant value C_v .

With the help of the familiar formulae,

$$R = kN$$

and

$$Ne = 9658 \cdot 3 \cdot 10^{10} \text{ E. S. U. ;}$$

the constant k of Planck's formula was used to determine the Avogadro number and the charge of the electron e .

In spite of these successes the Planck hypothesis was looked upon with considerable suspicion for several years afterwards. It appeared more as a mathematical trick than as a fundamental physical truth. There was an objection to the quantum theory of radiation on the score that unidirectional quanta could not explain interference of waves. That the oscillators do not satisfy the law of equipartition of energy was also considered to be an objectionable feature. A serious difficulty was pointed out by Lorentz in (1910) who showed that in a weak field of radiation an appreciable time would be needed for the absorption of a quantum. On the other hand, it was assumed in Planck's proof that the time of absorption is negligible.

In view of these difficulties Planck himself did not countenance Einstein's theory of light quanta in the beginning and his own original hypothesis he modified making the oscillators emit discontinuously but absorb continuously. It was shown that Planck's formula of heat radiation could be deduced even on the modified hypothesis. Two years later, that is, in 1913, Planck returned to the classical hypothesis that both emission and absorption are continuous. Later he reverted again to the second hypothesis.

It is no longer necessary to postulate either linear oscillators or unidirectional quanta of radiation to account for the observed facts of heat radiation or the photoelectric phenomena. Like the grin alone that has survived of the Cheshire cat the hypotheses of Planck and Einstein, (11) and (12), retain their importance as mathematical formulae but the mechanism which suggested them is no longer in fashion. In fact physical processes as they are understood to-day in the light of wave mechanics are much too subtle to follow any mechanical model satisfying classical laws. When we speak of waves and linear oscillators in the quantum theory of radiation the existence of such mechanism is not implied, but the terms are used because they are very suggestive descriptions of certain mathematical formulae and equations.

2. EINSTEIN'S RELATIONS

Einstein's relation ³ for light quanta,

$$E = h\nu \quad (12)$$

should be studied along with the vector relation, also given by him for light,

$$\hat{p} = h \hat{\tau} \quad (19)$$

where \hat{p} is the momentum vector and $\hat{\tau}$ the wave-number vector. These equations bring out the correspondence between the terms of corpuscular representation such as E and \hat{p} and those of the wave representation, ν and $\hat{\tau}$. To understand the correspondence consider the relativistic vector for a particle,

$$mv_x, mv_y, mv_z, imc \quad (= i E/c)$$

where
$$m = m_0 / \left(1 - \frac{v_x^2 + v_y^2 + v_z^2}{c^2} \right)^{\frac{1}{2}}, \quad (20)$$

and m_0 is the rest mass. Similarly, there is the vector

$$k_x, k_y, k_z, i\nu/c \quad (21)$$

associated with the plane wave

$$\psi = \psi_0 \cos 2\pi (x k_x + y k_y + z k_z - \nu t). \quad (22)$$

It is known that a Lorentz transformation leaves

$$m (x v_x + y v_y + z v_z - E t/m) \quad (23)$$

invariant. Hence the phase of the plane wave can remain unchanged under a Lorentz transformation in form and value if the two four-vectors satisfy

$$\frac{k_x}{mv_x} = \frac{k_y}{mv_y} = \frac{k_z}{mv_z} = \frac{\nu}{E} = \frac{1}{h}, \quad (24)$$

where h is a function of the invariants,

$$k_x^2 + k_y^2 + k_z^2 - \frac{\nu^2}{c^2} \quad \text{and} \quad m^2 (v_x^2 + v_y^2 + v_z^2 - E^2/m^2 c^2).$$

For light both the invariants are zero (see I 98) and h may be taken as a constant, Planck's constant, in particular. Putting

$$\hat{\tau} = (k_x, k_y, k_z), \quad \hat{p} = m (v_x, v_y, v_z) \quad (25)$$

we have the two relations stated in the beginning. They are assumed to be true for corpuscles of matter as well. On this assumption was based the original work of de Broglie. An important contribution by Einstein to physics was the derivation of (12) and (19) by reconciling the particle and wave representations of light; the extrapolation of these results for small particles of matter was done later by de Broglie.

SUPPLEMENT

In case the invariance of (23) is not obvious the following straightforward proof⁴ may be read.

For a Lorentz transformation,

$$\begin{aligned}x' &= \frac{x - ut}{(1 - u^2/c^2)^{1/2}}, \quad y' = y, \quad z' = z, \quad t' = \frac{t - ux/c^2}{(1 - u^2/c^2)^{1/2}}, \\ \dot{x}' &= \frac{v_x - u}{1 - uv_x/c^2}, \quad \dot{y}' = \frac{v_y(1 - u^2/c^2)^{1/2}}{1 - uv_x/c^2}, \quad \dot{z}' = \frac{v_z(1 - u^2/c^2)^{1/2}}{1 - uv_x/c^2}\end{aligned}$$

where $\dot{x}' = dx'/dt$, $v_x = dx/dt$, etc. Also $v^2 = v_x^2 + v_y^2 + v_z^2$, and similarly \dot{v} is defined. It may be verified that

$$\dot{x}'\dot{v}_x + \dot{y}'\dot{v}_y + \dot{z}'\dot{v}_z - c^2t' = \frac{(xv_x + yv_y + zv_z - c^2t)(1 - u^2/c^2)^{1/2}}{(1 - uv_x/c^2)}$$

and that

$$(1 - v'^2/c^2) = \frac{(1 - u^2/c^2)(1 - v^2/c^2)}{(1 - uv_x/c^2)^2}.$$

Hence

$$\begin{aligned}& \frac{m_0}{(1 - v'^2/c^2)^{1/2}} [\dot{x}'\dot{v}_x + \dot{y}'\dot{v}_y + \dot{z}'\dot{v}_z - c^2t'] \\ &= \frac{m_0}{(1 - v^2/c^2)^{1/2}} [xv_x + yv_y + zv_z - c^2t].\end{aligned}$$

3. THE SUCCESS OF THE QUANTUM HYPOTHESIS

In discussing the failure of classical mechanics we have described three difficulties which the theory could not successfully combat. As we have already seen Planck's hypothesis explains perfectly the observed rise and fall of the intensity of black-body radiation with wave length. It was natural to expect that the same hypothesis might account for the statistical situation⁵ also. For the resistance⁶ of metals and for the heat of rotation of diatomic molecules either the law of distribution of energy with temperature as given by Planck's theory was taken as true or the assumption $\oint p dq = nh$ which Planck made in the particular case of an oscillator, was used to regulate a more complicated mechanism. A certain measure of success was attained in most of these cases but nowhere was it as complete as in the problem of heat radiation. It looked as if a good clue was obtained in Planck's hypothesis but that more clues were needed before statistical properties of atoms and molecules could be clearly understood.

The third difficulty, caused by the remarkable stability of atoms and molecules, was brought nearer home by the new experimental findings. Already a model of the atom was proposed by J. J. Thomson according to which the positive charge is distributed in a sphere of radius 10^{-8} cm., and the negative charges which are electrons are elastically bound to it in positions of stable equilibrium. When the atom is slightly disturbed negative charges execute small oscillations which were supposed to explain the spectra. However this failed to provide an explanation of the simple formula of Balmer lines of hydrogen, viz.,

$$\nu = R \left(\frac{1}{2^2} - \frac{1}{n^2} \right)$$

where ν is the frequency, R the Rydberg constant and n , a positive integer exceeding 2. Nor was the Stark effect on spectral lines in an electrical field, discovered during the early formative period of Bohr's theory, understood by means of the Thomson model. Moreover, Rutherford's experiments on the scattering of α particles showed very pronounced deviations in their orbits which could be attributed only to the supposed concentration of positive charges in the centre of the atom. The experiments being of a decisive nature the Thomson model was replaced by Rutherford's solar-system model. The idea of a positively charged nucleus with a number of satellite electrons goes back to Rutherford. Although the scattering experiments strongly favoured this model there were serious objections against it. For instance the dynamical stability of a satellite electron is not possible if the classical theory of radiation due to moving charges is true. At this stage Bohr entered upon the scene and although he could not meet this particular objection he showed how Rutherford's model, modified by the quantum hypothesis, can explain the hydrogen spectra, the Stark effect and several other phenomena. Incidentally it may be observed that Rutherford's scattering formula provided a stimulus in the right direction, in 1913, when Bohr's theory was in the offing and an effective test in 1926, for verifying Bohr's probability interpretation of Schrödinger's wave function.

REFERENCES

1. REICHE, F., *The Quantum Theory*, pp. 1-30 (1930).
 2. BOHR, N., *Atomic Theory and the Description of Nature*, p. 94 (1934).
 3. FRENKEL, J., *Wave Mechanics*, Vol. I, pp. 15-16 (1936).
 4. MILNE, E. A., *Relativity, Gravitation and World Structure*, p. 89 (1935).
 5. REICHE, F., *Loc. Cit.*, pp. 61-86 (1930).
-

III FIRST QUANTUM THEORIES OF MATTER AND RADIATION

1. Postulates and Hypotheses of Bohr's Theory. 2. Successes of Bohr's Theory.
3. Einstein's Theory of Radiation. 4. Failures of the Two Theories.

1. POSTULATES AND HYPOTHESES OF BOHR'S THEORY

Bohr's theory, as it was first advanced in 1913, had to be considerably elaborated by Sommerfeld and others to meet with the growing demands of spectral evidence. Let it be clearly understood therefore that it is the postulates and hypotheses of the more elaborate theory, which had grown to its full stature in 1923, that will be considered here. They¹ are as follows :

- (A) An atom can exist in any one of a possible set of stationary states, each state having its own constant of energy.
- (B) When from a state of higher energy E_2 an atom falls to a state of lower energy E_1 a quantum of radiation is emitted with frequency ν , given by

$$h\nu = E_2 - E_1. \quad (1)$$

Similarly, an atom rises to a state of higher energy with the absorption of a quantum of radiation, and the increase in energy equals the energy of the quantum.

- (C) The laws of classical mechanics apply to the atom. But in the case of conditionally periodic systems the action variables satisfy the quantum conditions of Wilson and Sommerfeld,

$$\oint p dq \equiv I = nh, \quad (2)$$

n being a non-negative integer. There are as many quantum conditions as there are independent co-ordinates. For degenerate systems the reduced number of independent action variables satisfy the analogous conditions

$$J = nh. \quad (3)$$

- (D) The moving changes in the atom obey the laws of classical electrodynamics. The state of polarization of an emitted quantum of radiation and the question whether a transition is possible or forbidden are decided by Bohr's correspondence principle.

Every atom is distinguished by its atomic number which is some positive integer m ; that means it has a positive charge me in the nucleus and m electrons, each with a negative charge $-e$, rotating about it. Thus H has the atomic number 1, He has 2, Li has 3 and so on. The original classification of the elements of the periodic table was made by Mendeléeff on the basis of atomic weights but Moseley's conclusive studies of atomic spectra made it abundantly clear in 1913 that the periodic properties are dependent upon the atomic number. These facts were used in developing the theory of the atom on the assumptions stated above.

There is enough evidence that atoms exist in stationary states. That a stationary state can be characterized simultaneously by absence of radiation and moving charges is amply borne out by paramagnetic atoms like bismuth and ferromagnetic atoms like iron, cobalt and nickel. The apparent paradox is now resolved by the assumption of electronic spin.

In elucidating (B) some important considerations have to be presented. When the energy of the atom is the least possible it is said to be in the ground state; any other possible state is an excited state. As will be seen later two or more states may correspond to the same energy. The various states are usually characterized by different electronic orbits. Certain transitions from one level to another are highly improbable and they are called 'forbidden'. The theoretically forbidden lines may be detected with finer instruments. Whenever this has happened it has given a further proof of the accuracy of wave mechanics, these smaller and 'neglected' terms of calculation accounting for the fainter and 'forbidden' lines. It is possible, however, for an atom to change from one excited state to another, both characterized by the same energy, the distinguishing feature being different electronic orbits in the two states. Such a transition is radiationless and is described as Auger effect. What actually happens is that a normal atom exposed to X-rays gets into an excited state from which it sinks to a state of lower energy by the emission of a quantum. This quantum is further absorbed by one of the outer electrons usually making its orbit hyperbolic; and moving in the hyperbolic orbit the electron speeds off to infinity leaving the atom ionized. Thus the transition from one state to another is marked by an absence of radiation.

The quantum numbers are introduced by (2) and (3). For conditionally periodic systems the action variables I have been proved to be adiabatic invariants. So the quantum numbers also are adiabatic invariants of such a system. Hence the same quantum numbers can be used for describing the various excited states of hydrogen, helium or any other atom. The adiabatic invariants of a system are certain unchanging structural relations which exist even when the system changes—very slowly, no doubt—through the variation of certain physical parameters. The quantum numbers, therefore, play the role of the tensor invariants of general relativity in the present theory. Any con-

ditionally periodic system can be shown to be obtainable from a system of harmonic oscillators by a slow variation of certain parameters. Thus the Wilson-Sommerfeld conditions (3) are equivalent to Planck's hypothesis for an oscillator. In the case of a degenerate system the smaller number of independent action variables J are shown by Burgers to be adiabatic invariants.

It has been already shown that for mechanical systems such as those considered by Bohr

$$\delta E = \Sigma \omega \delta I. \quad (4)$$

For a transition from one set of large quantum numbers n to another set n' such that $n \sim n'$ is small compared to n and n' we can write (4) in the form

$$\delta E = h \Sigma (n' - n) \omega,$$

or, using (1)

$$\nu = \Sigma (n' - n) \omega. \quad (5)$$

Hence for large quantum numbers the Bohr frequencies are the same as those given by the classical theory. But (5) would not be true for small quantum numbers as in such a case δI would be comparable to I if expressed as τh where τ is an integer and ω cannot be considered to be unchanged for a finite variation in I . However, by Bohr's correspondence principle, (5) is assumed to be true for small quantum numbers as well. Secondly, only certain frequencies $\Sigma \tau \omega$ appear in the Fourier expansions of the co-ordinates of a mechanical system. The same frequencies ν are supposed to be allowed in Bohr's theory. Hence the transition from one level to another occurs only if $[n' - n]$ has the same set of values as τ in any of the frequencies. Thus arise the selection rules of Bohr's theory. In this manner the correspondence principle restricts the possible transitions between two levels, the allowed frequency of emitted or absorbed radiation being a possible frequency according to the classical theory. It also makes the quantum condition (1) equivalent to the classical relation (4) for large quantum numbers. The polarization of emitted radiation is determined by calculating the direction of the electric intensity vector as in the classical theory, the frequency ν being connected with the quantum numbers by (5).

At the common boundary of the microscopic and macroscopic domains the equations of classical mechanics must be equivalent to those of Bohr's theory. This is achieved by the correspondence principle. The discontinuity of quantum processes and the continuity of classical laws are apparently irreconcilable but the correspondence principle harmonises the two at the meeting ground of the two theories. Logic assigns an important place to this principle in any theory of atomic laws; and so we shall have to discuss it again when we consider Schrödinger's and Dirac's theories and also certain questions of radiation.

2. SUCCESSES OF BOHR'S THEORY

The outstanding successes⁸ of Bohr's theory will now be briefly considered. For our purpose the following topics need be taken up :

- (1) The spectra of hydrogen-like atoms.
- (2) The Stark effect.
- (3) Collision experiments of Franck and Hertz.

By a hydrogen-like atom is meant one with a nucleus of mass M and positive charge Ze , there being one satellite electron of mass m and charge $-e$. Neutral hydrogen, singly ionized helium, He^+ and doubly ionized lithium Li^{++} are examples of such an atom. For hydrogen, $M/m = 1845$, approximately. M being in general several thousand times as great as m the motion of the nucleus may be neglected in the first instance. Assuming the Coulomb law of electrostatic attraction we have for the kinetic energy and the potential,

$$T = \frac{1}{2} m (\dot{r}^2 + r^2 \dot{\theta}^2), \quad V = -\frac{Ze^2}{r}$$

$$p_1 \equiv \frac{\partial T}{\partial \dot{r}} = m\dot{r}, \quad p_2 \equiv \frac{\partial T}{\partial \dot{\theta}} = mr^2\dot{\theta}.$$

Hence the Hamiltonian is

$$H \equiv T + V = \frac{1}{2m} \left[p_1^2 + \frac{1}{r^2} p_2^2 \right] - \frac{Ze^2}{r}. \quad (6)$$

If the constant of energy is α_1 the Hamilton-Jacobi equation becomes

$$\frac{1}{2m} \left[\left(\frac{\partial F}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial F}{\partial \theta} \right)^2 \right] - \frac{Ze^2}{r} = \alpha_1. \quad (7)$$

Putting $p_2 = \frac{\partial F}{\partial \theta} = \alpha_2$

we get $\frac{\partial F}{\partial r} = p_1 = \left[2m \left(\alpha_1 + \frac{Ze^2}{r} \right) - \frac{\alpha_2^2}{r^2} \right]^{1/2}.$

Hence the quantum conditions are

$$I_2 = \oint p_2 d\theta = 2\pi \alpha_2 = n_2 h, \quad (8)$$

$$I_1 = \oint p_1 dr = \oint dr \left[2m \left(\alpha_1 + \frac{Ze^2}{r} \right) - \frac{\alpha_2^2}{r^2} \right]^{1/2} = n_1 h. \quad (9)$$

Eliminating α_2 between the last two equations we get for the total energy,

$$E = \alpha_1 = -\frac{2\pi^2 m Z^2 e^4}{n^2 h^2}. \quad (10)$$

where $n = n_1 + n_2. \quad (11)$

As we have shown earlier this is a case of degeneracy with only one action invariant $J = I_1 + I_2$. The frequency of emitted radiation is given by

$$\nu = RcZ^2 \left(\frac{1}{s^2} - \frac{1}{n^2} \right) \quad (12)$$

$$\text{where} \quad R = \frac{2\pi^2 e^4 m}{h^3 c}, \quad (13)$$

and s and n are positive integers so that $n > s$. R is the Rydberg constant whose value agrees with the empirical estimate of the same. For $Z=1$, $s=1$ gives the Lyman frequencies and $s=2$, the Balmer series.

If the motion of the nucleus is not ignored the value of the Rydberg constant becomes

$$\frac{2\pi^2 e^4}{h^3 c} \cdot \frac{m}{\left(1 + \frac{m}{M}\right)} \quad (14)$$

As M increases with the atomic number the constant increases in value. This is in perfect agreement with the results of spectroscopic observations.

A closer examination of the hydrogen spectra by a spectroscope of high resolving power however reveals a fine structure which could not be explained by the simple formula (12). Sommerfeld showed that if we start with the relativistic Hamiltonian, in place of (6), a formula for the fine structure is obtained.

He started with the Lagrangian

$$L \equiv T - V = mc^2 \left[1 - \left(1 - \frac{\dot{r}^2 + r^2 \dot{\theta}^2}{c^2} \right)^{1/2} \right] + \frac{Ze^2}{r}, \quad (15)$$

$$\text{getting} \quad p_1 = \frac{\partial L}{\partial \dot{r}} = \frac{m\dot{r}}{(1 - \beta^2)^{1/2}}, \quad p_2 = \frac{\partial L}{\partial \dot{\theta}} = \frac{mr^2 \dot{\theta}}{(1 - \beta^2)^{1/2}},$$

$$\text{where} \quad \beta^2 = (\dot{r}^2 + r^2 \dot{\theta}^2)/c^2.$$

The Hamiltonian is

$$H \equiv \dot{r} p_1 + r \dot{\theta} p_2 - L \equiv c \left(m^2 c^2 + p_1^2 + \frac{p_2^2}{r^2} \right)^{1/2} - mc^2 - \frac{Ze^2}{r}. \quad (16)$$

Substituting $\partial F/\partial r$ for p_1 and $\partial F/\partial \theta$ for p_2 in the energy integral,

$$H = \alpha_1, \quad (17)$$

we have the Hamilton-Jacobi equation

$$\left(\frac{\partial F}{\partial r} \right)^2 + \frac{1}{r^2} \left(\frac{\partial F}{\partial \theta} \right)^2 = \frac{1}{c^2} \left(\alpha_1 + \frac{Ze^2}{r} \right) \left(\alpha_1 + \frac{Ze^2}{r} + 2mc^2 \right). \quad (18)$$

$$\text{We can take} \quad \frac{\partial F}{\partial \theta} = \alpha_2, \quad (19)$$

so that

$$L = \oint \frac{\partial F}{\partial r} dr = \frac{2\pi Ze^2 (m + \alpha_1/c^2)}{[-\alpha_1 (2m + \alpha_1/c^2)]^{1/2}} - 2\pi (\alpha_2^2 - Z^2 e^4/c^2)^{1/2}, \quad (20)$$

$$I_2 = \oint \frac{\partial F}{\partial \theta} d\theta = 2\pi \alpha_2. \quad (21)$$

Eliminating α_2 between (20) and (21) we get the energy E :

$$E = \alpha_1 = -\frac{mc^2 \mu^2 Z^2}{2(I_1 + I_2)^2} \left\{ 1 + \frac{h^2 \mu^2 Z^2}{(I_1 + I_2)^2} \left(\frac{1}{4} + \frac{I_1}{I_2} \right) \right\} \quad (22)$$

where

$$\mu = 2\pi e^2/ch.$$

μ is the so-called fine-structure constant and its value is approximately 1/137.

The quantum conditions

$$I_1 = n_1 h, \quad I_2 = n_2 h,$$

give

$$E = -\frac{mc^2 \mu^2 Z^2}{2(m_1 + m_2)^2} \left\{ 1 + \frac{\mu^2 Z^2}{(n_1 + n_2)^2} \left(\frac{1}{4} + \frac{n_1}{n_2} \right) \right\}. \quad (23)$$

It is clear from the expression (22) that this is not a degenerate system and that two independent quantum numbers characterize each energy level. Writing in the familiar notation,

$$n_1 + n_2 = n, \quad n_2 = k \quad (24)$$

and

$$2\pi^2 m e^4 / ch^3 = R, \quad (13)$$

$$E = -RhcZ^2 \left\{ \frac{1}{n^2} + \frac{\mu^2}{n^4} \left(\frac{n}{k} - \frac{3}{4} \right) \right\}. \quad (25)$$

Hence the line of frequency ν as given by (12) is replaced by a series of lines with frequency $\nu + \Delta\nu$ so that

$$\nu + \Delta\nu = RcZ^2 \left(\frac{1}{s^2} - \frac{1}{n^2} \right) + RhcZ^2 \mu^2 \left[\frac{1}{s^4} \left(\frac{s}{k'} - \frac{3}{4} \right) - \frac{1}{n^4} \left(\frac{n}{k} - \frac{3}{4} \right) \right]$$

$$\text{or} \quad \Delta\nu = RhcZ^2 \mu^2 \left[\frac{1}{s^4} \left(\frac{s}{k'} - \frac{3}{4} \right) - \frac{1}{n^4} \left(\frac{n}{k} - \frac{3}{4} \right) \right], \quad (26)$$

k and k' being the quantum numbers corresponding to I_2 in the two states. It looks as if there is a double infinity of lines of emission corresponding to a transition from the state n to the state s since k and k' are still free. But we know from (24) and from the result that the zero value of n_2 makes the orbit run through the origin that

$$0 < k \leq n, \quad 0 < k' \leq s. \quad (27)$$

There is a further restriction put on the relative values of k and k' by the correspondence principle. For we know that the frequency which is given by

$$(s - n) \omega_1 + (k' - k) \omega_2 \quad (28)$$

must be a possible frequency according to the classical theory. The following equations have now to be recalled:

$$W_1 = \frac{\partial F}{\partial I_1}, \quad W_2 = \frac{\partial F}{\partial I_2}. \quad (29)$$

From the form of F we find that in the expansions of the coordinates of the problem, x, y or r, θ frequencies only of the type

$$\tau\omega_1 + \omega_2 \quad (30)$$

enter. Hence the selection rules are derived by comparing (28) with (30); $|n - n'|$ can have any integral value while $|k - k'|$ can have only the value unity. Thus

$$k - k' = \pm 1. \quad (31)$$

The conditions (27) and (31) control the multiplicity of the values of $\Delta\nu$ as given by (26).

The problem of the Stark effect is concerned with the motion of an electron moving under the Coulomb law of attraction to the nucleus and a uniform field force. Neglecting the complications caused by the presence of other electrons and the small relativistic correction the Hamiltonian is deduced from

$$T + V \equiv \frac{1}{2}m [\dot{z}^2 + \dot{\rho}^2 + \rho^2\dot{\phi}^2] - \frac{Ze^2}{r} + F_0 ez, \quad (32)$$

where cylindrical co-ordinates ρ, ϕ, z are used and $r^2 = \rho^2 + z^2$. F_0 is the intensity and Oz the direction of the uniform field of force. When these co-ordinates are used our results for a conditionally periodic system cannot be applied; the adiabatic invariants which are necessary for quantizing the system do not exist. If, however, parabolic co-ordinates are used the quantum conditions can be set down and the theory applied to solve the problem. This naturally raised suspicion as to whether the quantum formulation of mechanical problems depended on the choice of the co-ordinate system. The discussion that followed was terminated by Einstein who showed in 1917 that the quantum conditions can be expressed in a form independent of the co-ordinate system chosen. When parabolic co-ordinates are used the expression for energy becomes

$$E = -\frac{2\pi^2 m Z^2 e^4}{(I_1 + I_2 + I_3)^2} - \frac{3F_0}{8\pi^2 m Z e} (I_1 - I_2) (I_1 + I_2 + I_3), \quad (33)$$

where I_1, I_2, I_3 are the action invariants. Putting $I_1 = n_1 h, I_2 = n_2 h, I_3 = n_3 h$, we get E in terms of the quantum numbers. As the system is degenerate we use the transformation,

$$\left. \begin{aligned} I_1 + I_2 + I_3 &= J_1 = nh, \\ I_2 - I_1 &= J_2 = kh, \end{aligned} \right\} \quad (34)$$

so that

$$E = -\frac{2\pi^2 m e^4 Z^2}{n^3 h^2} - \frac{3F_0 h^3 n k}{8\pi^2 m Z e}. \quad (35)$$

It is the correspondence principle that will have to be used to determine the restrictions on possible transitions. That means the solution of the Hamil-

ton-Jacobi equation, F , is expressed as a function of I_1 and I_2 . The angle-variables \dot{W}_1, \dot{W}_2 are then given by

$$\dot{W}_1 = \frac{\partial F}{\partial J_1}, \quad \dot{W}_2 = \frac{\partial F}{\partial J_2}. \quad (36)$$

$$\text{As} \quad \dot{W}_1 = \omega'_1 t + s_1, \quad \dot{W}_2 = \omega'_2 t + s_2, \quad (37)$$

it is possible to express $x + iy$ and z as expansions of the form

$$\sum A' e^{2\pi i (\tau_1 \omega'_1 + \tau_2 \omega'_2) t}. \quad (38)$$

The conditions for polarization parallel to Oz are given by the expansion for z . Those for circular polarization are found from the expansion for $x + iy$.

If the energy level is supposed to be characterized by the principal quantum number only, any transition from one value of it to another gives not merely one line, as the formula (35) shows, but more lines due to the second quantum number k for each of the two states. This phenomenon, described as the splitting of the n -s line, was found to be in complete accord with observation in 1917 and it inspired great confidence in the theory.

If an atom consists of many electrons the energy of the atom as a whole can be considered to be, by a mathematical artifice, the sum of the energies of the satellite electrons. The artifice consists in compensating the terms of electronic interaction. The quantum hypothesis implies the existence of discrete levels of electrons corresponding to different energy values. An atom becomes excited as one or more electrons of the ground state change their levels under an external stimulus which imparts just the energy requisite for excitation. This theoretical inference was quantitatively verified in a series of collision experiments carried out by Franck and Hertz in 1919 and later on, by other investigators also. Consider in its ground state or state of lowest energy an atom like H or Li, or Na in which there exists one outer electron, the other electrons, if any, being bound in closed shells or sub-shells to the nucleus. We can denote possible levels of increasing energy for this electron by positive integers. The least energy needed for the electron to change its level is the difference in energy of the two lowest levels between which a transition is probable. The emission for a transition from the higher level to the other gives the resonance line. Now Franck and Hertz found out that if electrons possess an energy less than a critical value E_c they suffer an elastic impact in a collision with an atom of the type described above. If, however, the electrons possess the energy E_c there is an inelastic impact, the total energy is communicated to the atom and its outer electron jumps on to the next higher level. The observed value of E_c agrees perfectly with the energy difference of the first two levels. Agreement was also reached with the theory in more complicated experiments dealing with other excited states of atoms obtained by atom-electron collisions.

In cataloguing the successes of Bohr's theory we have avoided any reference to the properties of the periodic table. It is true that the classification of the elements in the periodic table was done by means of four quantum numbers, in 1925, when wave mechanics was still unknown. But giving credit to Bohr's theory for this explanation of the periods of the elements will not be right. We will not discuss the quantum numbers and the periodic table here as this question has to be taken up again. However, this is just the place where one may record that many of the early successes of the theory turned out to be failures in reality, when the underlying phenomena were examined with greater instrumental precision and in the light of fresh evidence. As an example we may consider the explanation of the fine structure of the $H\alpha$ line of hydrogen as obtained on Sommerfeld's theory. The five components as spectroscopically determined were given by the theory but the selection rules derived from the correspondence principle did not permit all the five of them. Moreover, if the fine structure were due only to the difference in relativistic effects on the masses of the moving nucleus and electron the fine structure of Na could not be accounted for. As was seen later the explanation of the alkali doublet and of the five components of $H\alpha$ rests on another relativistic effect, the spin of the electron.

3. EINSTEIN'S THEORY OF RADIATION

Einstein's work on light quanta was hailed as a triumph of Newton's corpuscular theory in several orthodox circles. The quanta of radiation are however essentially distinct from Newton's corpuscles of light. The quanta behave as particles always moving with the velocity of light; their rest mass is zero and in free space their energy and momentum (scalar) are connected by the relation

$$E/c = p; \quad (39)$$

they do not obey the classical deterministic laws of motion. On the other hand, Newton's corpuscles moved with changing velocities in a field of force or in a varying medium, their rest mass was not zero and they were subject individually to the law of causation. Experiments have now definitely disclosed that the quanta of radiation also carry angular momentum with them. Thus a beam of light falling on a screen may set the latter in rotation through the angular momentum of photons. The corpuscles of the classical theory were not supposed to carry angular momentum individually in this manner.

The two quality relations on which Einstein's theory is based have been previously stated. They are

$$E = h\nu, \quad \hat{p} = h\hat{\tau}. \quad (40)$$

These relations of energy and momentum combined with the principles of conservation of momentum and energy explain the photoelectric phenomena and the Compton scattering which could not be understood on the classical

wave theory of light. If the energy and momentum were carried by waves, as was believed on the wave hypothesis, it appeared a mystery how even a faint beam of monochromatic radiation of sufficiently high frequency could pull out the outer electrons of an exposed metal when an intense beam only of a frequency less than a critical value completely failed to do so. Einstein succeeded in unravelling the mystery by means of the simple energy equation,

$$\frac{1}{2} m v^2 = h\nu - P, \quad (41)$$

where v is the velocity of the ejected electron of mass m , $h\nu$ is the energy of the quantum of radiation and P is the binding energy of the outer electron in the atom combined with the work done in pulling the electron against the surface force of the metal. If $h\nu < P$ no electron can be pulled out. The equation clearly shows that the effect is controlled not by the intensity of the beam but by its frequency. Not only metals but non-metals and even inert gases can have their electrons pulled out by radiation provided the frequency is high enough. For radiations of the X-ray range, that is, for wave lengths less than 100 \AA and exceeding 10^{-12} cm , the Compton effect is witnessed in their interaction with the inner or bound electron of the atom. Suppose that an electron at rest is imparted a velocity v by a quantum of frequency ν_0 and that as a result of the impact the scattered light has frequency ν . The conservation of energy is fulfilled by

$$h\nu_0 = h\nu + mc^2 \left[(1 - v^2/c^2)^{-\frac{1}{2}} - 1 \right], \quad (42)$$

while the conservation of momentum in two perpendicular directions in the plane of motion is contained in

$$\frac{h\nu_0}{c} = \frac{h\nu}{c} \cos \theta + mv (1 - v^2/c^2)^{-\frac{1}{2}} \cos \phi, \quad (43)$$

$$0 = -\frac{h\nu}{c} \sin \theta + mv (1 - v^2/c^2)^{-\frac{1}{2}} \sin \phi. \quad (44)$$

θ is obviously the angle of scattering. Eliminating v and ϕ we get

$$\nu = \frac{\nu_0}{(1 + 2\alpha \sin^2 \theta/2)}$$

where

$$\alpha = \frac{h\nu_0}{mc^2}.$$

The wave-length shift due to scattering is

$$\Delta\lambda = c \left(\frac{1}{\nu} - \frac{1}{\nu_0} \right) = \frac{h}{mc} (1 - \cos \theta).$$

The maximum scattering corresponds to the angle $\theta = \pi/2$, in which case it is h/mc . This term is known as the wave length of the Compton shift. Similarly, many other phenomena such as fluorescence and the Raman effect,

discovered in 1928, can be understood in terms of the light quanta in so far as only the energy and momentum properties of radiation are involved therein.

Einstein's theory was not developed so as to provide laws of space-time behaviour for the photons. In the early stages of Schrödinger's wave mechanics it was customary to speak of the probability of position for a photon. This concept was not fully worked out like the cognate one of the probability for an electron and it did not make any contribution to the understanding of radiation. Heitler² makes the remark: "But there is no indication that, for instance, the idea of the 'position of a light quantum'; (or the 'probability for the position') has any simple physical meaning".

Photons are indistinguishable and classical statistics cannot therefore be applied to them. We may apply either the Fermi-Dirac statistics or the Einstein-Bose statistics. Logically the two are on equal footing and it is not known why photons follow the Einstein-Bose statistics. Planck's law of black-body radiation cannot be deduced if the alternative statistics is assumed for the quanta of radiation.

This distinction between the three statistics can be indicated by means of a trivial example. Suppose that there are two atoms A and B and three places a, b, and c for them. The following different distributions are possible :

Classical Statistics :

A, a b c a b a c b c
B, a b c b a c a c b

Einstein-Bose Statistics :

A, a b c a b c
B, a b c b c a

Fermi-Dirac Statistics :

A, a b c
B, b c a

According to the classical statistics as A and B are distinguishable there are nine different arrangements possible. But if one cannot distinguish A from B only the six different arrangements of the Einstein-Bose statistics are observable. If, however, the Fermi-Dirac statistics is employed it means that there is a further bar on the indistinguishable atoms which prevents two of them from being in the same place. The respective distributions tell us that the probability of one atom being in a and another in b is $2/9$, $1/6$ or $1/3$ according to the nature of the statistics followed.

We will not indulge here in a derivation of Planck's formula based on the Einstein-Bose statistics. A property of the photons that was made use of in the classical derivation ought to be noted here. Just like the electron the photon has three coordinates to mark out its position in space. The electron

is further distinguishable by two possible spin values $\pm 1/2$. Similarly two perpendicular directions are associated with the photon to describe its state of polarization. Both the directions are perpendicular to the line of motion. Any other direction of polarization can then be described in terms of probability.

Einstein's theory was by no means a complete theory of radiation. It had to be supplemented by Bohr's rules of radiation deduced from classical electrodynamics and the correspondence principle. In spite of the many successes of Einstein's theory and of the correspondence principle it appeared impossible to reconcile the two theories with each other. There was also the difficulty of harmonising the corpuscular theory and the classical wave theory of light. We shall now proceed to state the major difficulties that these first theories of matter and radiation confronted and that they could not surmount.

4. FAILURES OF THE TWO THEORIES

Nature reveals some phenomena and provokes the scientist to find the logical key to them. "Guess" says Nature and the scientist puts forth an intelligent guess. "You are quite clever but not right, so guess again", comes the challenge of Nature who discloses some more secrets and the scientist guesses more accurately. And Nature smiles and nods patronizingly. More secrets are revealed and each time the scientist uses all the resources at his command to give a more accurate answer. This goes on till at last the scientist finds his progress arrested by natural limitations. In the progress of the modern theories of atomic physics one witnesses the game very vigorously played.

Einstein's theory of the quanta provoked two questions :

- (1) What is radiation?—waves or corpuscles?
- (2) Is radiation spherical in space or directed?

The first question arose because interference and diffraction phenomena could be explained only on the classical wave theory of radiation while the photoelectric and Compton effects definitely contradicted that theory. The theory of photons, on the other hand, accounted for the effects but proved to be definitely ineffective in the treatment of the other phenomena and simple results of the wave theory such as the formula for the resolving power of a telescope could not be derived in terms of photons. Briefly, the situation was this ; the kinematical effects were explained by the wave theory, the dynamical effects by the particle theory. A logical synthesis of the two was reached in the quantum theory of radiation where the dynamical effects are realized by the rules of quantization superposed on the classical theory giving the kinematical effects. One thing the examination of this particular difficulty brought to light before the advent of the more abstract theories of Heisenberg and Dirac, viz., that the love of mechanical models must be sacrificed to the

greater love of logic if the complicated phenomena of matter and radiation are to be understood as rationally interconnected.

According to the classical theory an electric oscillator radiates energy in spherical waves. At great distances from the oscillator the energy being spread over a large surface is very small per unit area. This goes against the observed result that the energy of radiation even at great distances is found to be $h\nu$ in interaction with matter. Considering the energy exchanges in a state of equilibrium between Bohr atoms and radiation Einstein has established the law of black-body radiation. But considerations of momentum exchanges have shown that the radiation emitted by the Bohr atom has the character of a bullet. This is certainly in agreement with the photon hypothesis. What is curious however is that the character of the radiation should be correctly describable by means of selection and polarization rules when the momentum and energy are governed by the photon theory. In answer to the second question regarding the nature of radiation one had to admit that it might consist of tiny sets of directed waves at times and at times of spherical waves as given by classical electrodynamics. This question brings out more fully the significance of our earlier remark that the kinematical effects are given by the wave theory although the dynamical ones call for the photon hypothesis. In the absence of a complete logical picture of the abstract processes of radiation our knowledge of the subject was in an extremely unsatisfactory state in 1925.

Bohr's theory of the atom was considered unsatisfactory in 1925 for the following reasons :

(1) Approximate solutions of the many-body problem showed a discrepancy between the observed energy levels and the calculated ones. The ionizing potential for He, for instance, was quite different from the calculated one. (If V is the ionizing potential— Ve is the energy of the outermost electron). Moreover, the electric orbits as given by classical mechanics are not dynamically stable.

(2) The existence of stationary states is in sharp contradiction to classical electrodynamics. An electron experiencing an acceleration loses energy by radiation. The continuous loss of energy would ultimately bring the electron to the nucleus. The main problem of the atom is to reconcile the existence of stationary states with moving charges inside and this remained untouched on account of the convenient but faulty assumption that electrons move in their quantum orbits without a loss of energy by radiation.

(3) There was no rule for calculating the intensity of emitted lines. The correspondence principle gave accurate results only in special cases.

(4) The problem of scattering of light by an atom presented very grave difficulties. An atom was treated as a multiply periodic system and it was supposed to be disturbed by radiation. The result of the disturbance on the scattered light was found to involve the orbital periods of the electrons. But the periods are not observable and hence the solution of the problem of scatter-

ing was partly couched in unobservable terms. In the study of the problem of dispersion in Bohr's theory, Heisenberg was very much impressed with the absolute necessity of keeping all unobservables out of the theory on account of the unwelcome appearance of orbital periods in the results calculated for observational verification.

(5) A number of difficulties had accumulated during 1913-25 which could only be removed later by the hypothesis of the electron spin. There was the question of half quantum numbers without which the multiplet structure and the molecular spectra could not be accounted for. The separation of components in the anomalous Zeeman effect as compared to the separation of the normal components introduced the mystery of the Londe'-g-factor. Theories of ferromagnetism, electrical conductivity and valency were too rudimentary yet.

To study all this from the standpoint of Bohr's theory is purely of historical interest. It will suffice to remark here that Bohr's theory had to be continually modified to accommodate new experimental facts during the twelve years preceding 1925 and it became increasingly clear about this time (i) that the methods were not competent enough to cope with some situations (ii) that they led to wrong results definitely in certain cases and (iii) that there was no hope of building a synthetic logical structure out of classical mechanics and electrodynamics, ad hoc quantum hypotheses and the correspondence principle. For all these reasons the theory ceased to provide an incentive and a lead to further experimental work.

More important conclusions followed from this than the one that Bohr's theory must be scrapped. The failures of science prove to be the earnest of a coming great success. The circumstances were favourable for a new outlook in physics, for the scientist to leap forth and hazard a new guess. Above everything else Bohr's theory revealed the inadequacy of mechanical models in physical representations. The microscopic domain cannot be explored part by part so that in every part laws of classical mechanics hold good, the parts themselves being bound together by laws outside the pale of the classical theory.

The flames of thought that were to illumine the field and remove these obscurities were provided by de Broglie, Schrödinger, Born, Heisenberg and Dirac.

REFERENCES

- (1a) RUARK AND UREY, "*Atoms, Molecules and Quanta*", pp. 14, 65, 86, 168-80 (1930).
- (1b) REICHE, F., "*The Quantum Theory*," pp. 93, 108 (1930).
- (1c) CONDON AND MORSE, "*The Quantum Mechanics*" pp. 1-10 (1930).
- (2) HEITLER, W., "*The Quantum Theory of Radiation*", p. 64 (1936).
- (3) BIRTWISTLE, G., "*The Quantum Theory of the Atom*", pp. 223-29 (1926).
- (4) FRENKEL, J., "*Wave Mechanics*," Part I, pp. 12-13, 150 (1936).

IV WAVES AND PARTICLES

1. A Fundamental Analogy. 2. The Schrödinger Equation. 3. The two Interpretations of ψ . 4. Some Collision Phenomena. 5. The Kepler Problem.

1. A FUNDAMENTAL ANALOGY

The main function of analogies is to suggest an inner logical form. There is no harm in using them in scientific expression on the definite understanding that they prove nothing and that the attendant particulars are irrelevant. In the following is set forth, in detail, a fundamental analogy between geometrical optics and classical mechanics which has been responsible for the Schrödinger equation and a new philosophy of matter. First we expound the connection between the classical wave equation of physical optics and Fermat's principle on which geometrical optics is based. Next it is shown that there is a common symbolic expression for Fermat's principle and for the principle of least action which is equivalent to the laws of classical mechanics so far as conservative systems are concerned. The most ingenious step is the third which establishes a connection between classical mechanics and Schrödinger's equation, of the same logical form, which connects geometrical optics with the wave equation of light.

Consider¹ the propagation of an optical phenomenon which is represented by the function $u(x, y, z, t)$. If $V(x, y, z)$ is the velocity of propagation, c/V is the refractive index and the wave equation,

$$\nabla^2 u - \frac{1}{V^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad (1)$$

is satisfied by u . It is possible to express u in the form

$$u(x, y, z, t) = A(x, y, z) \cos 2\pi\nu [t - \psi(x, y, z)] \quad (2)$$

where ν is a constant. At any instant the surfaces of constant phase are

$$\psi(x, y, z) = \text{const.} \quad (3)$$

The curves which are orthogonal to this family of surfaces are called rays. The phase velocity is the velocity with which one has to travel along a ray in order to keep to a given value of the phase, $t - \psi$. Hence if V' is the phase velocity,

$$V' \delta t = \delta \psi, \quad (4)$$

where the components of δr are related to $\delta \psi$ by the condition of the constancy of phase, viz.,

$$\delta t = \frac{\partial \phi}{\partial x} \delta x + \frac{\partial \phi}{\partial y} \delta y + \frac{\partial \phi}{\partial z} \delta z; \quad (5)$$

since

$$\delta x = r \delta r = \delta r \frac{\partial \phi}{\partial x} / \left\{ \left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 + \left(\frac{\partial \phi}{\partial z} \right)^2 \right\}^{1/2} \quad (6)$$

and δy and δz are similarly related to δr ,

$$V' = \left\{ \left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 + \left(\frac{\partial \phi}{\partial z} \right)^2 \right\}^{-1/2} = \left| \text{grad } \phi \right|^{-1}. \quad (7)$$

In general, V and V' are not equal. But the two are equal for the wave forms of geometrical optics given by

$$u = C \cos 2\pi \nu \left[t - \int \frac{dr}{V} \right], \quad (8)$$

C being a constant. Wave optics reduces to geometrical optics when the condition

$$V = V' \quad (9)$$

is fulfilled. From the form of u in (8) and from the definition of a ray it follows that rays satisfy

$$\delta \int \frac{dr}{V} = 0, \quad (10)$$

which is Fermat's principle. Let us now express the condition (9) in quite a different form for a future application. The solution (2) is the real part of

$$A e^{\frac{2\pi i \nu t}{e} - \frac{2\pi i \phi}{e}} \quad (11)$$

where

$$\phi = -\psi \nu. \quad (12)$$

The expression (11) satisfies the wave equation provided

$$-4\pi^2 |\text{grad } \phi|^2 + 2\pi i \nabla^2 \phi + 4\pi^2 \nu^2 / V^2 = 0. \quad (13)$$

Hence if $\nabla^2 \phi$ is negligible compared to $|\text{grad } \phi|^2$ we have from (7), (12) and (13) that $V = V'$. Therefore, (9) is equivalent to the condition that

$$\nabla^2 \phi / |\text{grad } \phi|^2 \text{ is negligible.} \quad (14)$$

If ν is the frequency, $\lambda = V/\nu$ is the wavelength. On differentiating ϕ in an arbitrary direction l ,

$$\begin{aligned} \frac{\partial \phi}{\partial l} &= -\nu \frac{\partial \psi}{\partial r} \frac{dr}{dl} = -\nu |\text{grad } \psi| \cos \theta = -\frac{\nu}{V} \cos \theta, \\ \frac{\partial^2 \phi}{\partial l^2} &= \nu \frac{\cos \theta}{V^2} \frac{dV}{dl} = \frac{\cos \theta}{\lambda V} \frac{dV}{dl}. \end{aligned}$$

Hence the condition (14) is equivalent to the statement that

$$\left(\frac{1}{\lambda V} \frac{dV}{dt} \right) / \left(\frac{1}{\lambda} \right)^2 \text{ is negligible,}$$

$$\text{or that } \lambda \frac{dV}{V dt} \text{ is negligible,} \quad (15)$$

for any arbitrary direction. This means that the change in the velocity over a distance of the order of the wave length is negligible if the laws of geometrical optics are to be valid. Here is the reason why the interference and diffraction phenomena cannot be explained by means of geometrical optics.

Assuming that Einstein's equations for light,

$$E = h\nu \text{ and } p = h\tau = h/\lambda = h\nu/V \quad (16)$$

are valid also for material particles an infinitesimal change in the optical phase becomes

$$2\pi\nu \left[dt - \frac{dr}{V} \right] = \frac{2\pi}{h} [Edt - pdr]. \quad (17)$$

It follows that Fermat's principle is equivalent to the principle of least action,

$$\delta \int p dr = 0. \quad (18)$$

It was argued, therefore, by de Broglie that just as geometrical optics fails when the condition (15) breaks down so also classical mechanics becomes invalid in a domain of atomic dimensions where the phenomena must be controlled by some equation analogous to (1). By (16)

$$V = h\nu/p, \quad (19)$$

and so making this substitution in (1) the wave equation of mechanics was first obtained as

$$\nabla^2 u - \frac{h^2 \nu^2}{p^2} \frac{\partial^2 u}{\partial t^2} = 0. \quad (20)$$

In relativistic mechanics, the equation of energy in its usual notation for a particle is

$$E = \frac{m_0 c^2}{(1 - v^2/c^2)^{1/2}} + V(x, y, z) \quad (21)$$

and the momentum p is given by

$$p = \frac{m_0 v}{(1 - v^2/c^2)^{1/2}} = \frac{1}{c} [(E - V)^2 - m_0^2 c^4]^{1/2}. \quad (22)$$

Similarly, in Newtonian mechanics

$$E = \frac{1}{2} m_0 v^2 + V(x, y, z), \quad (21')$$

$$p = m_0 v = \sqrt{2m_0 (E - V)}. \quad (22')$$

Hence the wave equation is

$$\nabla^2 u + \frac{4\pi^2}{h^2 c^2} [(E - V)^2 - m_0^2 c^4] u = 0, \quad (23)$$

or
$$\nabla^2 u + \frac{8\pi^2}{h^2} m_0 (E - V) u = 0, \quad (23')$$

according as the mechanics is relativistic or Newtonian.

Let us suppose that a solution of either of these wave equations is of the form (11) with

$$\phi(x, y, z) = \frac{1}{h} S(x, y, z). \quad (24)$$

If the condition (14) holds good S satisfies

$$|\text{grad } S|^2 = \frac{(E - V)^2}{c^2} - m_0^2 c^4, \quad (25)$$

or
$$|\text{grad } S|^2 = 2m_0 (E - V). \quad (25')$$

S therefore plays the role of action in the Hamilton-Jacobi equation. Thus when the condition (14) is satisfied the wave equation reduces to a crucial equation of classical mechanics. In view of the fact that $\nabla^2 \phi$ is of the order of $1/h$ while $|\text{grad } \phi|^2$ is of the order of $1/h^2$ the condition (14) is equivalent to the statement that

$$h \text{ is negligible;} \quad (26)$$

classical mechanics must be replaced by the more accurate wave theory when h is not negligible.

It was Schrödinger who first pointed out that, since $\lambda \pm h/p$ for atomic particles, λ is of atomic dimensions if the experimental values of h and p are substituted. Hence, by (15), in regions of atomic dimensions the wave theory would be valid but not classical mechanics.

We can partially understand now how the new theory provides the representation of a particle in terms of waves. The optical analogy suggests the identification of the group velocity with the particle velocity. For a particle with velocity v , moving in free space,

$$E = \frac{m_0 c^2}{(1 - v^2/c^2)^{1/2}}, \quad p = \frac{m_0 v}{(1 - v^2/c^2)^{1/2}}$$

and so

$$v dp = d'E.$$

Hence, as $p = h\tau$ and $E = h\nu$

$$v = \frac{dE}{dp} = \frac{d\nu}{d\tau}. \quad (27)$$

Thus v turns out to be the group velocity. We shall consider presently the physical significance of this purely formal result.

We have also

$$v/\tau = E/p = c^2/v$$

which is the wave velocity. As $v < c$, the wave velocity definitely exceeds c and the waves of de Broglie cannot therefore be considered as physically real.

In the case of light, in ether,

$$E = pc$$

and so the wave velocity is

$$v/\tau = E/p = c$$

while the group velocity is also

$$dv/d\tau = dE/dp = c.$$

The two velocities are identical for radiation.

2. THE SCHRÖDINGER EQUATION

The non-relativistic classical Hamiltonian of a particle is given in the usual notation as

$$H \equiv \frac{1}{2m_0} [p_x^2 + p_y^2 + p_z^2] + V = E. \quad (28)$$

If we substitute for p_x , p_y and p_z the operators

$$p_x = -\frac{i\hbar}{2\pi} \partial/\partial x, \quad p_y = -\frac{i\hbar}{2\pi} \partial/\partial y, \quad p_z = -\frac{i\hbar}{2\pi} \partial/\partial z, \quad (29)$$

the energy integral leads to the Schrödinger equation (23') which may now be written as

$$\nabla^2 \psi + \frac{8\pi^2 m_0}{\hbar^2} (E - V) \psi = 0, \quad (30)$$

where ψ is the Schrödinger function. This suggested the postulate that for any conservative classical system for which

$$H(q, p) = E, \quad (31)$$

the corresponding Schrödinger equation is

$$H(q, -\frac{i\hbar}{2\pi} \partial/\partial q) \psi = E \psi. \quad (32)$$

The correspondence is generally responsible for a certain ambiguity if Cartesian coordinates are not used. The point may be made clear by considering a particular term p_r^2/r^2 which appears in the Hamiltonian when polar coordinates are used. The question now is whether this term is to be replaced by

$$\frac{1}{r^2} \frac{\partial^2}{\partial r^2} \quad \text{or} \quad \frac{\partial^2}{\partial r^2} \frac{1}{r^2}. \quad (33)$$

The Hamiltonian of the wave theory does not therefore follow unambiguously from the classical theory if any arbitrary system of coordinates is used. We have pointed out earlier that a similar situation had arisen in the treatment of the Stark effect on Bohr's theory as the Wilson-Sommerfeld conditions could not be expressed in terms of invariants in a perfectly general coordinate system. The present question was tackled by Podlosky in 1928 with a view to getting the wave Hamiltonian unambiguously when the coordinates are not Cartesian.

For non-conservative systems the total energy, when it exists, is a function of time; so

$$E = \frac{i\hbar}{2\pi} \partial/\partial t \quad (34)$$

gives the operator to be used for E in such systems. (29) and (34) may be looked upon as postulates of Schrödinger's theory. But it is possible to deduce them from the postulates of Dirac's theory of states and observables. We have already shown the connection between Poisson brackets and non-commutative relations of the type,

$$qp - pq = \frac{i\hbar}{2\pi}. \quad (35)$$

It is sufficient to remark here that (35) is satisfied by

$$p = -\frac{i\hbar}{2\pi} \partial/\partial q. \quad (36)$$

For if $f(q)$ is any arbitrary differentiable function of q

$$(qp - pq)f(q) = \frac{i\hbar}{2\pi} [f(q) + qf'(q)] - \frac{i\hbar}{2\pi} qf'(q) = \frac{i\hbar}{2\pi} f(q). \quad (37)$$

The unusual substitutions (29) and (34) are thus connected with the Poisson brackets of the classical theory. In Dirac's theory it is not the substitutions that are fundamental but relations such as (35). Given a function $\phi(p)$ instead of $f(q)$ we may take

$$q = \frac{i\hbar}{2\pi} \partial/\partial p, \quad (38)$$

so that
$$(qp - pq)\phi(p) = \frac{i\hbar}{2\pi} \phi(p). \quad (39)$$

The substitutions of the form (38) are excluded in Schrödinger's theory but they are permitted by the more general non-commutative relations of Dirac's theory. The relation (38) is used to get a wave equation similar to Schrödinger's for the momentum probabilities of a particle. In addition to the significance of differential operators what may be emphasized here is the greater generality of Dirac's theory.

It is interesting to see how the Wilson-Sommerfeld conditions are incorporated in the theory of Schrödinger's equation². The substitution of the

operation as given by (34) extends the scope of the theory beyond the multi-periodic systems of Bohr's theory. Already we have seen that

$$\log \psi = \frac{2\pi i S}{h} + 2\pi i \nu t, \quad (40)$$

which is true for any conservative system of particles. Hence one of the adiabatic invariants for a conditionally periodic system is

$$\oint p dq = \oint \frac{\partial S}{\partial q} dq = \frac{h}{2\pi i} \oint \frac{\partial}{\partial q} (\log \psi) dq;$$

$\log \psi$ is a multiple-valued function so that we get

$$\oint p dq = nh, \quad (41)$$

where n is zero or an integer. The value of n will naturally depend upon the form of ψ as a function of q .

In the early stages of wave mechanics de Broglie gave an interesting explanation of the quantum numbers. If in a string with two ends fixed stationary vibrations are set up there is always an integral number of nodes on the join of the fixed ends. If $n+1$ is the number of nodes, d the distance between consecutive nodes and D the length of the join, $nd = D$. On this analogy it was argued that an electron in a Bohr orbit represents a stationary vibration of n nodes so that

$$\int \frac{ds}{\lambda} = n.$$

But when we use the condition, $h/\lambda = p = mv$,

$$\int mv ds = nh, \quad (42)$$

which is equivalent to a quantum condition in Bohr's theory. In fact this was the first empirical success of de Broglie's bold assumption that the two quality relations, which Einstein had deduced for light, are also true for material corpuscles. The later experiments on the diffraction of electron beams made by G. P. Thomson, Rupp and others showed that the quality relations may be assumed to be valid even for atomic particles participating in aperiodic phenomena.

Some remarks may be made regarding the mathematical nature of ψ . In general ψ will be complex as otherwise the interpretation of ψ functions in terms of observables and probabilities completely fails. Along with the partial derivations that appear in Schrödinger's equations ψ must be single-valued and finite so that the integral of $|\psi|^2$ over the configuration space is finite and constant in time. How the last condition affects the complexion of the problem is most easily seen in the illustration of the harmonic oscillator. We have for an oscillator,

$$H = \frac{1}{2m} p^2 + \frac{1}{2} \mu x^2 = E, \quad (43)$$

giving the Schrödinger equation,

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2m}{h^2}(E - \frac{1}{2}\mu x^2)\psi = 0,$$

which may be expressed as

$$\frac{d^2\psi}{dx^2} + (\lambda - \alpha^2 x^2)\psi = 0, \quad (44)$$

where $8\pi^2mE/h^2 = \lambda$ and $4\pi^2m\mu/h^2 = \alpha^2$. The only solution which satisfies the integral condition, n being a non-negative integer, requires

$$h = (2n + 1)\alpha, \quad (45)$$

$$\text{which means} \quad E = (n + \frac{1}{2})h\nu \quad (46)$$

$$\text{where} \quad \mu = 4\pi^2\nu^2m.$$

The form of ψ is given by

$$\psi_n = N_n e^{-\xi^2/2} H_n(\xi) \quad (47)$$

where N_n is a constant, H_n the Hermite polynomial of the n th degree and $\xi = \sqrt{\alpha}x$. The value of the constant is determined usually by the normalization condition, viz.,

$$\int_{-\infty}^{\infty} |\psi_n|^2 dx = 1. \quad (48)$$

When (48) is satisfied ψ_n is said to be normalized. The quantum number arises in this case through the integral condition which rejects solutions except those of the form (47). By Bohr's theory we have in place of (46)

$$nh = \oint p dx = \oint p \dot{x} dt = \oint 2T dt = \oint (T + V) dt = E \cdot 1/\nu.$$

The energy $\frac{1}{2}h\nu$ is called the residual or zero-point energy possessed by atoms near the absolute zero of temperature. It is customary to reduce a mathematical problem of atoms and molecules at the zero point to one of harmonic oscillators and then apply the above result of residual energy. This is how an attempt is made to explain some of the Van der Waals forces. For electronic distributions the zero-point energy is explained by means of the Fermi-Dirac statistics. Serious difficulties are still encountered regarding the behaviour of matter near the absolute zero; among the most notorious are the suprafluidity of helium and the superconductivity of metals. It is doubtful if the residual energy can account for them.

When the reader is introduced to the uncertainty principle it will be understood that, even in the state of lowest energy, an oscillating particle cannot be at once at rest and at the origin as that would violate the principle. Hence, even in the lowest state, the uncertainty of position and velocity is responsible for some energy which is the zero-point energy. In the theory

of radiation the analogy is provided by the transverse self-energy of the electron in free space.

3. THE TWO INTERPRETATIONS OF ψ

Schrödinger's first theory² of ψ is known as the hydrodynamical theory. Consider an electron of mass m and charge e moving in an external field of potential V so that ψ satisfies the equation,

$$\frac{\hbar^2}{8\pi^2 m} \nabla^2 \psi + eV\psi - \frac{\hbar}{2\pi i} \frac{\partial \psi}{\partial t} = 0. \quad (49)$$

It is usual to denote the conjugate complex function obtained by changing the sign of i in ψ as ψ^* . Evidently ψ^* satisfies the equation,

$$\frac{\hbar^2}{8\pi^2 m} \nabla^2 \psi^* + eV\psi^* + \frac{\hbar}{2\pi i} \frac{\partial \psi^*}{\partial t} = 0, \quad (50)$$

which is implicit in (49) and which is obtained therefrom by changing the sign of i throughout. The charge density ρ , the current density \hat{I} and the energy density u are defined by

$$\rho = -e\psi^*\psi, \quad (51a)$$

$$I = \frac{-eh}{4\pi im} (\psi^* \text{grad } \psi - \psi \text{grad } \psi^*), \quad (51b)$$

$$u = \frac{\hbar^2}{8\pi^2 m} \text{grad } \psi^* \cdot \text{grad } \psi. \quad (51c)$$

With these definitions the laws of conservation of charge, momentum and energy follow in the form,

$$\frac{d}{dt} \int \rho d\tau = 0, \quad (52a)$$

$$\frac{d}{dt} \int \hat{I} d\tau = -e \int \nabla V \psi^* \psi d\tau, \quad (52b)$$

$$\frac{d}{dt} \int u d\tau = \int eV \frac{\partial}{\partial t} (\psi^* \psi) d\tau, \quad (52c)$$

where $d\tau$ is the volume element and the integrations are carried over all space. To deduce (52a) and (52c) multiply (49) by ψ^* and (50) by ψ ; adding and subtracting the two gives the results when Green's theorem is applied. For (52b) multiply (49) by $\partial \psi^* / \partial x$, differentiate (50) with respect to x and then multiply it by ψ ; subtract one from the other and integrate the result using Green's theorem. The normalization condition of ψ makes it necessary that ψ should vanish at infinity. This fact has to be used in each of the above deductions.

The above interpretation of ψ was responsible for a wrong formula for

the intensity of emitted radiation. In the case of an atom as a conservative system the wave equation is

$$H\psi = E\psi \quad (53)$$

and, corresponding to the different permissible functions ψ_n which satisfy the equation, there are found to be discrete values of energy E_n . Thus E_n is the eigenvalue or characteristic value of E corresponding to the characteristic state ψ_n . If the reader is puzzled as to how these discrete values of energy and the discrete functions ψ_n arise the equation (53) may be likened to that of an ellipse. Corresponding to each direction there is a definite value of the central radius vector just as there is a value of E corresponding to each ψ . But only two mutually perpendicular directions, viz., those of the principal axes and the corresponding values of the vectors have a special significance inasmuch as the whole ellipse can be constructed with a knowledge of them. Similarly, of physical significance are the eigenfunctions ψ_n and the values E_n which give the full content of (53). Incidentally, it may be remarked that the eigenvalues may form a continuous range. As in the case of the harmonic oscillator both ψ_n and E_n follow from the integral condition on ψ . Now a large assembly of atoms⁵ is described by

$$\psi = \sum c_r \psi_r \quad (54)$$

where $|c_r|$ is the number of atoms in the state ψ_r . We may express ψ_r in the form

$$\psi_r = \psi_r^0(x, y, z) e^{-2\pi i E_r t/h} \quad (55)$$

For on substituting for ψ_r in

$$H\psi_r = \frac{ih}{2\pi} \frac{\partial \psi_r}{\partial t} \quad (56)$$

we get

$$H\psi_r = E_r \psi_r, \quad (57)$$

$$\psi^\times \psi = \sum \sum c_r \bar{c}_s \psi_r^0 \psi_s^0 e^{2\pi i (E_s - E_r) t/h}, \quad (58)$$

where \bar{c}_s is the conjugate complex of c_s . An absorption or emission between the levels E_r and E_s thus depends upon the number of atoms in both the states according to the first theory of ψ . This result contradicts experience. Even according to Bohr's theory the intensity of emitted radiation depends upon the number of atoms in the higher level only. The suggestion that the intensity of radiation can be determined from a formula such as (58) has failed.

Another failure⁶ of this interpretation of ψ , was occasioned by the supposition that electrons are charge clouds. If we assume that the different parts of a charge cloud repel each other we get wrong results definitely. More-

over, charge clouds will not behave as corpuscles in collision experiments. In a problem of two particles the charge cloud must be supposed to be spread over a space of six dimensions. These difficulties were enough to discredit the first interpretation of ψ although it gave a satisfactory explanation of the stationary states through those terms of (58) which do not change with time. It was Born who supplied the second interpretation and prepared the ground for Heisenberg's uncertainty principle.

According to Born, if ψ is the Schrödinger function for a system of n coordinates $q_1 q_2 \dots q_n$,

$$\psi^* \psi dq_1 dq_2 \dots dq_n \quad (59)$$

is the probability that, at time t , the system is in a configuration bound by q_r and $q_r + dq_r$ for $r = 1, 2, \dots n$. The first successful application of this interpretation was made for deducing the scattering formula for α particles obtained by Rutherford from classical principles. A simple example may be taken now introducing the new meaning of ψ .

For a particle moving along Ox in free space the wave equation runs as follows :

$$\frac{h^2}{8\pi^2 m} \cdot \frac{\partial^2 \psi}{\partial x^2} = \frac{h}{2\pi i} \frac{\partial \psi}{\partial t} \quad (60)$$

$$\text{A solution is} \quad \psi = A e^{2\pi i (\tau x - \nu t)} \quad (61)$$

where A is a constant and

$$-\frac{1}{2} \frac{h^2}{m} \tau^2 + h\nu = 0, \quad (62)$$

which is obvious when expressed in the usual notation as

$$\frac{1}{2} p^2/m - E = 0. \quad (62)$$

A particle of momentum p and energy E is expressed by the wave function (61). There is no such thing as a definite position associated with it. A particle whose dynamical qualities p and E are precisely defined is indefinite as regards its kinematical qualities inasmuch as they are represented by the infinite wave. An echo of Zeno's paradox of the arrow comes to the mind as one notices that velocity or momentum at an instant cannot be correlated with position.

Consider now a wave group defined by the equation,

$$\psi(x, t) = \int_{-\infty}^{\infty} a(\tau) e^{2\pi i (\tau x - \nu t)} d\tau \quad (63)$$

which satisfies (60), τ and ν being connected by (62). Let

$$a(\tau) = e^{-(\tau - \tau_0)^2/q^2} \quad (64)$$

so that τ_0 is the value corresponding to the maximum of $a(\tau)$. $2q$ defines the length of the symmetrical zone about the line of symmetry so that the ordinate falls here from the maximum to e^{-1} of it. q is taken as the measure of precision. The smaller the value of q the greater is the precision. Put

$$\tau = \tau_0 + q\xi \quad (65)$$

and

$$\begin{aligned} v &= \frac{1}{2} h \frac{\tau^2}{m} = \frac{1}{2} \frac{h}{m} (\tau_0 + q\xi)^2 \\ &= v_0 + \frac{1}{2} \frac{h}{m} (2\tau_0 q\xi + q^2 \xi^2). \end{aligned} \quad (66)$$

Hence

$$\begin{aligned} \psi(x, t) &= q e^{2\pi i (\tau_0 x - v_0 t)} \int_{-\infty}^{\infty} e^{-\xi^2 + 2\pi i (qx \xi - \frac{h\tau_0 q \xi t}{m} - \frac{hq^2 \xi^2 t}{2m})} d\xi \\ &= \sqrt{\frac{\pi}{\alpha}} q e^{-\beta^2/\alpha} e^{2\pi i (\tau_0 x - v_0 t)} \end{aligned} \quad (67)$$

where $\alpha = 1 + \frac{\pi h}{m} i q^2 t$, $\beta = \pi q \left(x - \frac{h\tau_0 t}{m} \right) = \pi q (x - v_0 t)$. (68)

Here v_0 is the velocity corresponding to the maximum coordinate of the wave packet. It is the average velocity associated with the particle. The probability distribution is given by

$$\psi^* \psi = \frac{\pi q^2}{\left\{ 1 + \left(\frac{\pi h q^2 t}{m} \right)^2 \right\}^{\frac{1}{2}}} \cdot e^{\frac{-2\pi^2 q^2 (x - v_0 t)^2}{1 + (\pi h q^2 t/m)^2}}. \quad (69)$$

The precision parameter now becomes

$$\frac{1}{\sqrt{2\pi q}} \left\{ 1 + \left(\frac{\pi h q^2 t}{m} \right)^2 \right\}^{\frac{1}{2}}. \quad (70)$$

As the parameter increases in value the wave group becomes more and more diffused. A particle moving with an average velocity v_0 in free space can be represented in wave mechanics by a Gaussian curve which spreads with time so that the maximum moves uniformly with velocity v_0 . It is obvious from (65) and (66) why we take $v_0/\tau_0 = v_0$ for the average velocity of the particle, which is by definition the average value of $dv/d\tau$.

Three conclusions are suggested by the above calculations. Firstly a solitary wave like (61) describes a possible state in which a particle can exist but tells nothing of the configuration. For the configuration we have to take a wave group like (63) and see how it is controlled by the wave equation. It gives us a more diffuse configuration with the passage of time so that the uncertainty associated with the average value increases. Incidentally, it is rather unfortunate that q should be called the precision parameter since an

increase of it means an increase of uncertainty. But what is equally important is that if we know the configuration at $t = 0$ the uncertainty increases not only for $t > 0$ but for $t < 0$ as well : this is the second point we should emphasize drawing attention to the form (70). Thirdly, configurations necessarily imply wave groups and hence with the same particle, in a particular configuration, different values of energy and momentum are associated through the concept of probability. How this is done may be explained with the help of some fundamental results.

Let ψ be the different possible normalized eigen- ψ 's of Schrödinger's equation,

$$H\psi = \frac{ih}{2\pi} \frac{\partial \psi}{\partial t} \equiv E\psi. \quad (71)$$

Then

$$\int \psi_r^* \psi_r d\tau = 1 \quad (71a)$$

and

$$\int \psi_r^* \psi_s d\tau = 0 \quad (71b)$$

the element $d\tau$ standing for $dq_1 \dots dq_n$ when q_1, q_2, \dots, q_n are the co-ordinates ; the integrations are carried over the whole q -space. The first result is merely a restatement of the normalization condition. The second implies the orthogonality of wave functions which is true if ψ_r and ψ_s correspond to different values of energy. For

$$\begin{aligned} H\psi_r^* &= E_r^* \psi_r^*, \\ H\psi_s &= E_s \psi_s. \end{aligned}$$

Hence

$$\psi_r^* H\psi_s - \psi_s H\psi_r^* = (E_s - E_r) \psi_r^* \psi_s. \quad (72)$$

The integration gives, on using the operational form of H and Green's theorem,

$$(E_s - E_r) \int \psi_r^* \psi_s d\tau = 0. \quad (73)$$

A more elegant proof is given on the basis of Dirac's theory. The orthogonality condition does not necessarily hold good if $E_r = E_s$, that is, if ψ_r, ψ_s are eigenfunctions of a degenerate state.

In general it is possible to express the eigen- ψ of a system by

$$\psi = \sum c_r \psi_r \quad (74)$$

or by an integral. c_1, c_2 etc. are in general complex numbers. If all the ψ 's are normalized

$$\sum |c_r|^2 = 1. \quad (75)$$

We now say that if the system is in the state ψ the probability that the energy of the system is E_r corresponding to ψ_r is $|\psi_r|^2$. Similarly, in the particle problem discussed above the resolution of $\psi(x, t)$ into wave functions of the form (61) gives the probabilities of different momentum values.

Are the probability waves⁸ real? No! They are just a logical fiction and their purpose is to explain the observed atomic phenomena. In a general problem the waves are multi-dimensional; the wave packets usually disintegrate rapidly; an actual discovery of a particle makes a wave-packet suddenly shrink in size. These are characteristic properties of wave packets which show that they belong essentially to the logical structure underlying the phenomena.

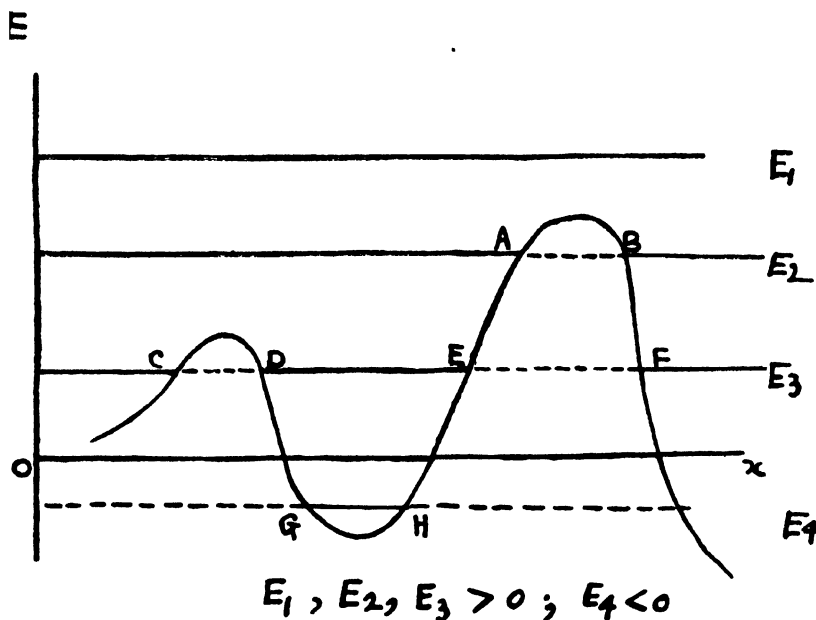
In classical mechanics a problem is solved when the coordinates are obtained as functions of time. In the new theory of wave mechanics the solution must answer adequately all possible questions of the form: if a dynamical observable has a value lying between u and $u + du$ what is the probability that another observable has a value lying between v and $v + dv$? The Schrödinger theory, as modified by Born's interpretation, did not provide answers to all such questions. The theory of ψ answered only the question: if the energy has a particular value what is the probability that the position configuration stands in a certain region of the coordinate space? The momentum configuration could not be deduced from the ψ -function. All these questions were later successfully attacked by Dirac's transformation theory.

The concept of probability has been most successfully applied to problems of collision and aperiodic phenomena in general. But at the stage where the atomic theory was brought in 1927 the phenomena of radiation were not satisfactorily explained. The relation between photons and waves and that between electromagnetic waves⁹ and probability were misunderstood or remained a mystery. A proper theory of the co-efficient of spontaneous emission and the electron spin was not yet deduced from the postulates and physical principles of the theory.

Schrödinger's theory in terms of probabilities was a definite advance on Bohr's theory. It gave a logical explanation of the three quantum numbers n, l, m . A successful derivation of the scattering formulae was also provided by the postulates of the theory. Although it did not explain the dual nature of light a new line of thought was initiated, independent of mechanical models and concerned with observed effects only. The new phenomenological outlook must be distinguished from the Aristotelian philosophy: the former proceeds from the phenomena to the logical correlation, the latter from the postulates to the phenomena. The change in attitude has been responsible for a new concept of particles which is quite different from the classical. It is not surprising therefore that the classical concept of causality finds no place in the scheme of new particles.

4. SOME COLLISION PHENOMENA

We will first describe certain phenomena* and then try to explain them by the method of the Fourier expansion and the probability postulate.



. Fig .

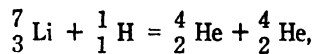
Consider an atomic particle moving in free space in a direction parallel to Ox . Suppose that a potential barrier is set up on the way as in the figure. The behaviour of the particle as a result of the barrier may first be described on the classical theory. Since

$$T + V = E$$

all motion for which $E < V$ is impossible. For the kinetic energy cannot be negative. There are four cases to be discussed :

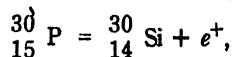
- (i) A particle moving with total energy E_1 is undisturbed by the barrier.
- (ii) If the total energy is E_2 a particle can move towards the barrier from either side. The subsequent motion will be influenced by the nature of impact at A or B. In any case a particle will not be able to jump over the peak.

- (iii) When the total energy is E_3 motion is possible in DE, to the right of F, or to the left of C. But no particle moving in any one of these zones can migrate to another.
- (iv) Only in the zone GH is motion possible with total energy E_4 . These four situations are often envisaged in atomic physics. We will now see in what way a particle of wave mechanics behaves in similar situations.
- (i') If the E_1 level is not very much above the potential peak and if the potential disturbance extends over a few de Broglie wave lengths the particle may be reflected back. When cathode rays are passed through a parallel-plate condenser it is discovered that some of the electrons are reflected back when the plates are brought sufficiently near together.
- (ii') If the E_2 level is not very much below the peak the so called "tunnel effect" is witnessed: the particle moving on the left has a finite probability of appearing on the right and vice versa. This effect is particularly marked in nuclear phenomena. Gamow's original theory of α decay was based on this effect. The equation of transmutation,



implies that a proton can penetrate through the barrier of repelling charges in the lithium nucleus and split it into two α particles. The experiments in which surface electrons are pulled out of metals by applying a uniform electric field outside provide a similar situation.

- (iii') A particle coming from the left of C with energy E_3 has a finite chance of going into a state of periodic motion in the region DE provided CD is equal to an integral number of de Broglie wave lengths. Similarly, if EF is also equal to an integral number of de Broglie wave lengths, the particle may appear on the right of F. This requirement is partially fulfilled in the theory of Ramsauer's scattering experiments on inert gases and in Oppenheimer's experiments on hydrogen atoms in an electric field.
- (iv') According to Dirac's theory the positron makes its appearance whenever an electron in one of the negative energy levels ascends to a level of positive energy on account of interaction with radiation or otherwise. In the transmutation,



an electron appears in the P nucleus from one of the negative-energy levels. This is similar to an electron of energy E_4 entering from outside into the region GH.

The explanation of all these phenomena lies in the possible Fourier expansions and the probability concept. By the well-known integral theorem of Fourier, a function $f(x)$ with a finite number of discontinuities and with its behaviour at infinity suitably controlled can be expressed as

$$f(x) = \int_{-\infty}^{\infty} A_{\tau} e^{2\pi i \tau x} d\tau$$

where

$$A_{\tau} = \int_{-\infty}^{\infty} f(x) e^{-2\pi i \tau x} dx.$$

Consider as a particular case

$$\begin{aligned} f(x) &= e^{2\pi i \tau_0 x} & \text{for } |x| < \frac{1}{2} \Delta x, \\ &= 0 & \text{for } |x| \geq \frac{1}{2} \Delta x, \end{aligned}$$

so that

$$A_{\tau} = \int_{-\frac{1}{2}\Delta x}^{\frac{1}{2}\Delta x} e^{2\pi i x (\tau_0 - \tau)} dx = \frac{\sin \pi (\tau_0 - \tau) \Delta x}{\pi (\tau_0 - \tau)}.$$

A_{τ} as a function of τ , has infinite maxima. The principal maximum corresponds to $\tau = \tau_0$ giving the value Δx . When

$$(\tau_0 - \tau) \Delta x = \pm 1$$

A_{τ} becomes zero and its value outside this zone of τ may be neglected as small. The smaller the value of Δx the greater becomes the significant zone about τ_0 . The usual diffraction formula,

$$d \sin \phi = n\lambda,$$

is based on a similar mathematical result. The smaller the breadth of the slit d , the greater is the significant extension of the diffraction pattern as measured by $\sin \phi$. Now an atomic particle moving with a definite energy and momentum in free space corresponds to a wave,

$$f(x, t) = e^{2\pi i (\tau x - \nu t)}.$$

The effect of a disturbance is to limit in some way the unlimited expansion of the wave. The partial expansion which is restricted by the disturbance can be expanded by the Fourier theorem into waves which correspond to different

values of τ and ν and hence of energy and momentum. The probabilities of these new states are given by the coefficients in the expansion according to (74) and (75). Every disturbance offered to an atomic particle increases its capacity to overcome it by providing new potentialities of state. If the disturbance is not sufficiently small calculations show that the effects described above do not occur.

It is necessary to prove that the probability theory of waves leads to the laws of classical mechanics at the common frontier of microscopic and macroscopic domains. For the other interpretation of ψ it has been already shown how the wave form reduces to the action function satisfying the Hamilton-Jacobi equation. Following Ehrenfest we take the simpler case of one particle moving along Ox under a force $-\frac{\partial V}{\partial x}$, V being the potential. The centre of gravity of the probability distribution is

$$X = \frac{\int \psi^* x \psi dx}{\int \psi^* \psi dx}$$

and if ψ is normalized

$$X = \int \psi^* x \psi dx.$$

Using the wave equation

$$\frac{h^2}{8\pi^2 m} \frac{\partial^2 \psi}{\partial x^2} - V\psi + \frac{ih}{2\pi} \frac{\partial \psi}{\partial t} = 0$$

we find

$$m\ddot{X} = \int \psi^* \left(-\frac{\partial V}{\partial x} \right) \psi dx.$$

The right-hand side gives the average value of $\partial V/\partial x$ for the distribution. Hence the Newtonian equation

$$m\ddot{X} = -\frac{\partial V}{\partial X}$$

follows. Thus the most probable distributions follow the classical laws. In the problem of a particle in free space we saw that the most probable position changes with time according to Newton's first law.

In the derivation of the strange results described in the beginning of this section where a de Broglie wave is partially reflected and partially transmitted the boundary conditions satisfied at the surface of disturbance are the classical ones, viz., (i) the continuity of ψ and (ii) the continuity of the normal derivative of ψ . These additional boundary conditions formally complete the theory of aperiodic phenomena to the limited extent of Schrödinger's method.

5. THE KEPLER PROBLEM

The most important application of Schrödinger's theory has been made to the hydrogen-like atom with a nuclear charge Ze and only one satellite electron. This is known as the Kepler problem on account of its resemblance to the problem of planetary motion. The electronic levels are determined in any atom, containing n electrons, by reducing the problem to the simpler one of a nucleus and one electron. Fortunately such a reduction is possible within the limits of approximation; otherwise the description of electronic motion in an atom would be very complicated. The theory of the periodic table and the application of the Pauli exclusion principle are based upon the assumption that, to the first degree of approximation, every electron orbit can be calculated in terms of the Kepler problem. The Schrödinger equation in this case runs as follows :

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2} \left(E + \frac{Ze^2}{r} \right) \psi = 0. \quad (76)$$

This is obtained on the assumption that the motion of the nucleus is negligible. We are interested in ψ -functions of the form $R(r)S(\theta)T(\phi)$ and hence we write the differential equation in polar coordinates,

$$\begin{aligned} \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial \psi}{\partial r} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) \\ + \frac{8\pi^2 m}{h^2} \left(E + \frac{Ze^2}{r} \right) \psi = 0. \end{aligned} \quad (76')$$

The equation splits up into the following three :

$$\frac{d^2 T}{d\phi^2} + \alpha T = 0, \quad (77)$$

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{dS}{d\theta} \right) - \frac{\alpha^2 S}{\sin \theta} + \beta S = 0, \quad (78)$$

$$\text{and } \frac{1}{r^2} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) - \frac{\beta R}{r^2} + \frac{8\pi^2 m}{h^2} \left\{ E + \frac{Ze^2}{r} \right\} R = 0. \quad (79)$$

To know how the splitting arises multiply (76') by $r^2 \sin^2 \theta / \psi$. Then we notice that the equation assumes the form

$$f(r, \theta) + f_1(\phi) = 0 \quad (80)$$

As r, θ, ϕ are independent coordinates (80) is equivalent to

$$f_1(\phi) = -\alpha, \quad f(r, \theta) = \alpha,$$

where α is some constant. Similarly the equation $f(r, \theta) = \alpha$ splits up into two, and the other constant β is introduced. Since ψ must be single-valued

and vanishing at infinity it is necessary that $\alpha = -m^2$, m being zero or an integer. Solving (77) we get the normalized functions.

$$T_m = \frac{1}{\sqrt{2\pi}} e^{im\phi}. \quad (81)$$

$$\int_0^{2\pi} T_m^* T_m d\phi = 1.$$

$$\text{For } m \neq n, \quad \int_0^{2\pi} T_m^* T_n d\phi = 0.$$

m is the magnetic quantum number. The azimuthal quantum number l is introduced in the definition of S . The restriction on the nature of ϕ permits only a polynomial in $\cos \theta$ and $\sin \theta$ as the solution of S . We get the Legendre polynomials.

$$S_{|m|}^l = P_{|m|}^l(\theta) \cdot \left[\frac{2l+1}{2} \cdot \frac{(l-|m|)!}{(l+|m|)!} \right]^{\frac{1}{2}} \quad (82)$$

where

$$\beta = l(l+1), \quad (83)$$

and l is a positive integer not less than $|m|$. The numerical factor of (82) is introduced so that the normalization condition

$$\int_0^\pi [S_{|m|}^l]^2 \sin \theta d\theta = 1$$

is satisfied. It may be verified that

$$\int_0^\pi S_{|m|}^l S_{|m|}^{l'} \sin \theta d\theta = 0, \quad l \neq l'.$$

Incidentally, we observe that between the azimuthal quantum number l and the magnetic quantum number m is the relation,

$$l \geq |m|. \quad (84)$$

If (84) is violated the S -functions do not exist. Similar considerations of (79) give the only possible solution

$$R_{nl}(r) = - \left[\left(\frac{8\pi^2 m Z e^2}{n h^2} \right)^{\frac{1}{2}} \cdot \frac{(n-l-1)!}{2n \{(n+l)!\}^{\frac{1}{2}}} \right]^{\frac{1}{2}} e^{-\rho/2} \rho^l L_{n-l-1}^{2l+1}(\rho) \quad (85)$$

where

$$\rho = \frac{8\pi^2 m Z e^2 r}{n h^2}$$

and
$$L_s^r = \frac{d^s}{d\rho^s} \left\{ e^\rho \frac{d^r}{d\rho^r} \left(\rho^r e^{-\rho} \right) \right\}.$$

The L -functions are the associated Laguerre polynomials. The integer n is such that

$$n \geq (l + 1)$$

and it is called the principal quantum number. The R_{nl} -functions are all normalized. The characteristic ψ -functions are now given by

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r) S_{|m|}^l(\theta) T_m(\phi). \quad (86)$$

They are orthogonal and normalized on account of

$$\int_0^\infty \int_0^\pi \int_0^{2\pi} |\psi_{nlm}|^2 r^2 \sin \theta d\theta d\phi dr = 1, \quad (87)$$

and

$$\int_0^\infty \int_0^\pi \int_0^{2\pi} \psi_{nlm}^* \psi_{n'l'm'} r^2 \sin \theta d\theta d\phi dr = 0, \quad (88)$$

where n, l, m and n', l', m' are not both the same set of quantum numbers. The energy constant assumes the eigenvalues,

$$E_n = - \frac{2\pi^2 m Z^2 e^4}{h^2 n^2}. \quad (89)$$

These are precisely the values obtained by Bohr using the quantum conditions. In wave mechanics the eigenvalues follow from the restrictions on the ψ -function of Schrödinger's equation.

The electronic levels for which $l = 0, 1, 2, 3, 4$ are called respectively s, p, d, f and g levels. The first four letters were suggested by the sharp, principal, diffuse and fundamental lines of the hydrogen spectra. $l = 4$ is called a g level because it follows the f level. The electrons for which $l = 0$ constitute the K shell. The higher values of n give the successive L, M, N, O etc. shells. Corresponding to a given value of l there are $(2l + 1)$ values of m given by (84). If the principal quantum number is n the different pairs of $l-m$ values are in number,

$$\sum_{l=0}^{n-1} (2l + 1) = n^2. \quad (90)$$

There is a fourth quantum number for an electron : it is the spin s which has the possible values $\pm \frac{1}{2}$. Hence in the n th shell the maximum number of electrons that can be packed is $2n^2$ if two or more electrons are not to have the same set of quantum numbers n, l, m, s . We are using here the exclusion

principle in giving each electron a separate wave function ψ_{nlms} . Thus in the K shell there can be only 2 electrons, in the L shell only 8, in the M shell only 18 and so on.

The connection between the energy constant and the quantum numbers merits attention. If we construct the Schrödinger equation from the classical Hamiltonian we get (89) in which case the energy depends only on the principal quantum number. If we take the relativistic Hamiltonian and solve the wave equation resulting therefrom the energy is found to be a function of both n and l . This is in good agreement with Sommerfeld's classical theory of fine structure. In a magnetic field the energy is influenced by the quantum numbers m and s also. Moreover the constant Z varies from atom to atom so that an electron in the $nlms$ level in one atom has a different energy from the electron in the same level in another atom, there is, one for which Z is different. In Bohr's theory, the adiabatic principle had to be invoked to justify the use of the same quantum numbers for different values of Z . No such difficulty arises in wave mechanics because of the manner in which the quantum numbers are introduced in the theory.

A physical interpretation of the quantum numbers l and m is provided by the differential operations of moments. The angular momentum of an electron about the z -axis is given by

$$\begin{aligned} H_z &= xmy - ymx \\ &= xP_y - yP_x \\ &= -\frac{ih}{2\pi} (x \partial/\partial y - y \partial/\partial x) = -\frac{ih}{2\pi} \partial/\partial \phi. \end{aligned} \quad (91)$$

When H_z operates on ψ_{nlm} we have from the form of T_m

$$H_z \psi_{nlm} = -\frac{ih}{2\pi} \cdot (im) \psi_{nlm} = \frac{hm}{2\pi} \psi_{nlm}. \quad (92)$$

The eigenvalues of the angular-momentum operator H_z are $hm/2\pi$. The resultant angular momentum is given by

$$\begin{aligned} H^2 &= H_x^2 + H_y^2 + H_z^2, \\ &= (yP_z - zP_y)^2 + (zP_x - xP_z)^2 + (xP_y - yP_x)^2, \\ &= -\left(\frac{h}{2\pi}\right)^2 \left[(y \partial/\partial z - z \partial/\partial y)^2 + (z \partial/\partial x - x \partial/\partial z)^2 + \right. \\ &\quad \left. (x \partial/\partial y - y \partial/\partial x)^2 \right] \\ &= -\left(\frac{h}{2\pi}\right)^2 \left[r^2 \nabla^2 - r \partial/\partial r \cdot r \partial/\partial r - r \partial/\partial r \right] \\ &= -\frac{h^2}{4\pi^2} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right]. \end{aligned}$$

Hence

$$\begin{aligned} H^2 \psi_{nlm} &= -\frac{\hbar^2}{4\pi^2} \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \right] \psi_{nlm} \\ &= \frac{\hbar^2}{4\pi^2} \beta \psi_{nlm} = l(l+1) \hbar^2 \psi_{nlm} / 4\pi^2. \end{aligned} \quad (9)$$

The expression for the eigenvalues of H is $\frac{\hbar}{2\pi} [l(l+1)]^{1/2}$. The connection

between the magnetic moment \hat{M} and the angular momentum \hat{H} is given on the basis of the correspondence principle by

$$\hat{M} = -\frac{e}{2m_0 c} \hat{H}. \quad (94)$$

Similarly if \hat{S} is the angular momentum due to the spin, we have the magnetic moment \hat{M}' given by

$$\hat{M}' = -\frac{e}{m_0 c} \hat{S}. \quad (95)$$

A rigorous derivation of this correspondence is extremely difficult. The hydrodynamical interpretation of ψ provides a derivation of (94). The other was derived by L. H. Thomas by considering the relativistic effect of the electronic motion on the precession as observed from the nucleus. Dirac's theory gives a rigorous explanation of both (94) and (95). It should also be noted that it is not right to associate a genuine spin with the electron in the mechanical sense. In (94) and (95) m_0 is the mass of the electron.

The associated Laguerre polynomials show that R_{nl} has $(n-l-1)$ zeros. The effect of it is that ψ vanishes on $(n-l-1)$ spheres. The function S has $(l-m)$ zeros while T_m has m only. Hence the total number of nodal surfaces for an electron vibrating in the state ψ_{nlm} is $(n-1)$. As $|\psi|^2$ is the probability density for the position of the electron at any time we notice that this is independent of time. The vibrations ψ_{nlm} of the electron are stationary. Hence there is no loss of energy due to radiation. We can interpret Bohr's quantum conditions as the conditions for stationary de Broglie waves. The nature of the function $|\psi|^2$ shows that there are n probability humps separated by the $(n-1)$ zeros. It is customary to refer to the probability distribution of the electron as the charge cloud as in Schrödinger's first theory of ψ .

The average value of any function f of the electronic co-ordinates is given by

$$\iiint \psi_{nlm}^* f \psi_{nlm} r^2 \sin \theta d\theta d\phi dr, \quad (96)$$

for the state ψ_{nlm} . The probability that the electron exists at a distance between r and $r + dr$, irrespective of the direction, is

$$\begin{aligned}
 & \int_0^{2\pi} \int_0^\pi \psi_{nlm}^* \psi_{nlm} r^2 dr \sin \theta d\theta d\phi \\
 &= R_{nl}^2 r^2 dr \int_0^{2\pi} \int_0^\pi |S_l^i|_m |T_m|^2 \sin \theta d\theta d\phi. \quad (97)
 \end{aligned}$$

We plot $R_{nl}^2 r^2$ against r and where the maximum is reached the analogue of the radius of Bohr's orbit is obtained.

Given a number of charges in a state of motion about the origin the field intensities and the flow of radiation can be determined as functions of the dipole and quadrupole moments of the charges by the laws of classical electrodynamics. For an atomic system with particles having charges e_1, e_2, \dots, e_k etc. the x -component of the dipole moment is

$$P_x^{nn'} = \sum e_k \int \bar{\psi}_n^* x_k \psi_{n'} d\tau$$

$$\text{where } \bar{\psi}_n = \psi_{nlm} e^{-2\pi i E_n t/\hbar}, \quad \bar{\psi}_{n'} = \psi_{n'l'm'} e^{-2\pi i E_{n'} t/\hbar}.$$

Hence the electric moment vector associated with the transition $n \rightarrow n'$ of an electron is

$$\hat{P}^{nn'} = \hat{R}^{nn'} e^{-2\pi i (E_n - E_{n'})t/\hbar} \quad (98)$$

$$\text{where } R_x^{nn'} = -e \int \psi_{nlm}^* x \psi_{n'l'm'} d\tau, \text{ etc. ;} \quad (99)$$

the intensities of the radiation are obtained from the vector $\hat{R}^{nn'}$ by using the classical formulae. It can be proved that if $|l - l'| \neq 0$, $R^{nn'} = 0$. Hence we obtain the selection rule,

$$\Delta l = \pm 1. \quad (100)$$

The system of charges may have no dipole moment as, for instance, a system consisting of two equal positive charges at two points and a negative double charge at the midpoint; the quadrupole moment must be considered in this case. We have now,

$$R_x^{nn'} = -e \int \psi_{nlm}^* x^2 \psi_{n'l'm'} d\tau. \quad (101)$$

The selection rule (100) no longer applies. Lines which are forbidden by the dipole moment are allowed by the formula (101). Like this effect there is another, of the second order, due to the magnetic dipole.

If there is a magnetic field along z let us take

$$x = \rho \cos \phi, \quad y = \rho \sin \phi, \quad d\tau = \rho d\phi dz d\rho,$$

$$\text{and } \psi_{nlm} = \chi e^{im\phi}, \quad \psi_{n'l'm'} = \chi' e^{im'\phi}$$

Hence
$$\int \psi_{nlm}^* x \psi_{n'l'm'} d\tau = \int \int \chi \chi' x \rho dz d\rho \int e^{i(m-m')\phi} d\phi. \quad (102)$$

The integral vanishes if

$$m - m' \neq 0. \quad (103)$$

Similarly calculating the other components,

$$\int \psi_{nlm}^* x \psi_{n'l'm'} d\tau, \quad \int \psi_{nlm}^* y \psi_{n'l'm'} d\tau \quad (104)$$

we find that they vanish unless

$$|m - m'| = 1. \quad (105)$$

The peculiarity of a dipole is that there is no emission of radiation in the direction of vibration. Hence when $m = m'$, as there is only the z -component of the dipole moment, there is no emission parallel to the field direction, Oz . The maximum flow of radiation is at right angles to Oz and it is polarized parallel to Oz according to a classical result of the oscillator. The radiation corresponding to $\Delta m = \pm 1$ is circularly polarized about Oz . An excellent brief account of this question is given in a general manner by Condon and Shortley. They have shown the exact relationship between the electric moment and the electric intensity which gives the direction of polarization.

An interesting result regarding $S'_{|m|}(\theta)$ and $T_m(\phi)$ is given by Unsöld. It is

$$\sum_{m=-l}^l \left[S'_{|m|}(\theta) \right]^2 \left[T_m(\phi) \right]^2 = \text{const.} \quad (106)$$

This theorem has important applications in the problems of many electrons.

The effect of a nuclear disintegration on the atomic levels may be mentioned in passing. Suppose for the sake of definiteness that an electron is emitted by the nucleus at $t = 0$ changing the effective nuclear charge from z to $z + 1$. Let the original wave functions be ϕ_1^z, ϕ_2^z , etc. and let the new ones be similarly indicated by $\phi_1^{z+1}, \phi_2^{z+1}$, etc. The characteristic value of the energy also will change. Consider an electron of one of the shells given by

$$\phi_1^z e^{-2\pi i E_1 t/h}. \quad (107)$$

As a result of the β emission, the possible states for the bound electron are ϕ_r^{z+1} for $r = 1, 2, 3, \dots$ etc. Hence,

$$\phi_1^z e^{-2\pi i E_1 t/h} = \sum C_r \phi_r^{z+1} e^{-2\pi i E'_r t/h} \quad (108)$$

at the time of emission, giving

$$\phi_1^z = \sum C_r \phi_r^{z+1}. \quad (109)$$

Therefore,
$$C_r = \int \phi_r^{z+1} \phi_1^z d\tau \quad (110)$$

The probability that the new energy is E' , is $|C_r|^2$. This also means that the probability that the outgoing electron takes away the energy $E - E'$ from the bound electron as it passes through the shell is $|C_r|^2$. Similar conclusions follow regarding the energy changes of the other bound electrons.

REFERENCES

- (1) DE BROGLIE AND BRILLOUIN, "*Wave Mechanics*", pp. 55-72, (1928).
- (2a) HEISENBERG, "*The Physical Principles of the Quantum Theory*", pp. 157-62 (1930).
- (2b) PAULING AND WILSON, "*Introduction to Quantum Mechanics*", pp. 50-90, pp. 112-50 (1935).
- (3) ZIMMER, E., "*The Revolution in Physics*", pp. 132-76 (1936).
- (4) CONDON AND MORSE, "*Quantum Mechanics*", pp. 208-31, pp. 87-91 (1930).
- (5) FRENKEL, J., "*Wave Mechanics*" Part I, pp. 41-6 (1936).
- (6) MOTT, N. F., "*An Outline of Wave Mechanics*", pp. 70-1 (1930).
- (7) BORN, M., "*Atomic Physics*", pp. 305-9 (1935).
- (8) HERZBERG, G., "*Atomic Spectra and Atomic Structure*", pp. 37-54, p. 105 (1937).

V MATRICES AND WAVES

1. Heisenberg's Matrix Theory. 2. The Harmonic Oscillator. 3. The Connection between Waves and Matrices. 4. Probability Postulates.

1. HEISENBERG'S MATRIX THEORY

From waves and corpuscles we have now to turn to matrices. The conspicuous failures of Bohr's theory had shown the utter futility of all attempts to visualize the atomic processes through pictorial representations. We have seen how Schrödinger's theory assumed a new and abstract form although it originated from de Broglie's attempt to reconcile two different representations of matter, as of radiation. Heisenberg's approach to the quantum theory was largely influenced by the difficulties that he and Kramers encountered in a treatment of dispersion by Bohr's theory. The main difficulty was that the conceptual element of the latter theory had grown out of proportion to do its job of correlating the empirical facts : and it inspired Heisenberg to initiate a new line of attack. Its essence lay in the phenomenological outlook, "from the surface to the centre" philosophy. The theory started from physical phenomena and ran on to the mathematical postulates.

For a conditionally periodic system the Fourier expansion of a coordinate runs as follows¹ :

$$q_k = \sum_{\tau_1 = -\infty}^{\infty} \sum_{\tau_2 = -\infty}^{\infty} \dots \sum_{\tau_f = -\infty}^{\infty} q_{\tau_1 \tau_2 \dots \tau_f}^{(f)} e^{2\pi i(\tau_1 \nu_1 + \tau_2 \nu_2 + \dots + \tau_f \nu_f)t} \quad (1)$$

The amplitudes $q_{\tau_1 \tau_2 \dots \tau_f}$ and the frequencies $\sum \tau \nu$ are functions of the "continuously variable constants" of integration. It is by varying the constants of integration that different dynamical states are obtained. The expansion (1) is not directly useful for a description of the quantum phenomena. Bohr had to supplement (1) by quantum rules. The phenomena are describable by terms of the form,

$$q_k(n_r, m_r) e^{2\pi i \nu(n_r, m_r)t} \quad (2)$$

where n_r and m_r are sets of integers each consisting of f . If the Bohr frequency condition is to hold good, there must be two sets of integers for distinguishing a transition between two states. Secondly, if the Rydberg-

Ritz combination principle is to hold good, the frequencies have to fulfil the law,

$$\nu(n_r, k_r) + \nu(k_r, m_r) = \nu(n_r, m_r). \quad (3)$$

Heisenberg's first problem was to establish correspondence between the terms (2) and the coefficients of the expansion (1). On the basis of classical electrodynamics it was assumed that terms of the form (2) describe fully the character of radiation and its intensity. Hence for a full-fledged calculus of observables such terms appeared sufficient. The correspondence principle for large quantum numbers, as it had been enunciated by Bohr, was used by Heisenberg to connect the terms (2) with those appearing in the classical expansion (1). One thing that looked quite clear at the outset was that the different terms connected with different sets of integers are independent of each other. The ensemble of discrete terms could therefore be arranged as a matrix,

$$Q_k(n_r, m_r) = \left\| q_k(n_r, m_r) e^{2\pi i \nu(n_r, m_r) t} \right\|, \quad (4)$$

the different rows being given by the different values of n_r and the columns by those of the other set of numbers. The matrices are assumed to be Hermitian so that

$$Q(n_r, m_r) = Q^*(m_r, n_r) \quad (5)$$

or as

$$\nu(n_r, m_r) = -\nu(m_r, n_r), \quad (6)$$

$$q_k(n_r, m_r) = q_k^*(m_r, n_r). \quad (7)$$

The last assumption was suggested by the presence of conjugate imaginaries in the expansion for q_k . Consider now how the correspondence is carried out for large quantum numbers between the classical terms and the matrix elements.

Using the conditions

$$I_k = n_k h, \quad \nu_k = \frac{\partial H}{\partial I_k} \quad (8)$$

of the classical theory we get, when n_k is large compared to α_k , that is $n_k \gg m_k$,

$$\nu_k = \frac{\partial H}{\partial I_k} = \frac{H(n_1, n_2, \dots, n_f) - H(m_1, m_2, \dots, m_f)}{\alpha_k h}$$

from which follows, for $n_r \gg (n_r - m_r)$,

$$\begin{aligned} \nu(n_1, n_2, \dots, n_f, m_1, m_2, \dots, m_f) &= \frac{1}{h} [H(n_r) - H(m_r)] \\ &\approx \frac{1}{h} [\Sigma(n_r - m_r) \frac{\partial H}{\partial n_r}] \\ &= \Sigma(n_r - m_r) \nu_r. \end{aligned} \quad (9)$$

In the Fourier expansions of q_k in the stationary state $H(n_r)$, there arises a term associated with the frequency $\sum (n_r - m_r) \nu_r$ and we identify that with $q_k(n_r, m_r)$. It must be emphasized that the matrix terms for small quantum numbers have no analogues in the classical theory.

For large values of n, m consider the (n, m) element of the matrix $(2\pi i/h)(qp - pq)$:

$$\begin{aligned} & \sum_k (2\pi i/h) [q(n, k) p(k, m) - p(n, k) q(k, m)] e^{2\pi i \nu(n, m) t} \\ &= \sum_{\alpha} \sum_{\beta} (2\pi i/h) [q(n, n - \alpha) p(n - \alpha, n - \alpha - \beta) - p(n, n - \beta) q(n - \beta, \\ & \qquad \qquad \qquad n - \alpha - \beta)] e^{2\pi i \nu(n, m) t}, \end{aligned}$$

where $\alpha + \beta = n + m$. Also

$$\nu(n, n - \alpha - \beta) = \sum_{r=1}^f (\alpha_r + \beta_r) \nu_r,$$

$$q(n, n - \alpha) - q(n - \beta, n - \beta - \alpha) \doteq \sum \beta_r \frac{\partial}{\partial I_r} q(\alpha_r, I_r),$$

$$\text{and } e^{2\pi i t \sum \beta_r \nu_r} p(n - \alpha, n - \alpha - \beta) \doteq e^{2\pi i t \sum \beta_r \nu_r} p(\beta, I).$$

$$\text{Hence } \frac{\partial}{\partial W_i} [p(\beta, I) e^{2\pi i t \sum \beta_r \nu_r}] = \beta_i \cdot 2\pi i p(\beta, I) e^{2\pi i t \sum \beta_r \nu_r}.$$

Using these results we get the (n, m) element of the matrix $\frac{2\pi i}{h}(qp - pq)$

$$\begin{aligned} &= \sum_{\alpha} \sum_{\beta} \left\{ [q(n, n - \alpha) - q(n - \beta, n - \beta - \alpha)] p(n - \alpha, n - \alpha - \beta) \right. \\ & \quad \left. - [p(n, n - \beta) - p(n - \alpha, n - \beta - \alpha)] q(n - \beta, n - \beta - \alpha) \right\} \\ & \qquad \qquad \qquad \frac{2\pi i}{h} e^{2\pi i \nu(n, n - \alpha - \beta) t} \\ &= \sum \left(\frac{\partial q}{\partial I_r} \frac{\partial p}{\partial W_r} - \frac{\partial p}{\partial I_r} \frac{\partial q}{\partial W_r} \right) \\ &= -[q, p]. \end{aligned} \tag{10}$$

The (n, m) element of the matrix turns out to be the $(n-m)$ harmonic of $-[q, p]$ expressed in terms of the action-angle variables. In view of this correspondence a postulate of Heisenberg's theory is stated as

$$(qp - pq) = \frac{i\hbar}{2\pi} [q, p] \tag{11}$$

where q and p are canonical conjugates of classical mechanics. From (11) and from the classical relations satisfied by the Poisson brackets we find Heisenberg's matrices subject to the following quantum conditions :

$$q_l p_l - p_l q_l = \frac{i\hbar}{2\pi} 1 \equiv \frac{i\hbar}{2\pi}, \quad (12a)$$

$$q_l p_k = p_k q_l, \quad (12b)$$

$$p_k p_l - p_l p_k = 0, \quad (12c)$$

$$q_k q_l - q_l q_k = 0, \quad (12d)$$

where k and l are distinct suffixes and by 1 in (12a) is meant a unit matrix of the same dimensions as the matrix on the left-hand side.

The Hamiltonian equations of the classical theory give

$$\dot{q} = [q, H], \quad \dot{p} = [p, H] \quad (13)$$

and hence it is postulated that,

$$\frac{i\hbar}{2\pi} \dot{q} = (qH - Hq), \quad \frac{i\hbar}{2\pi} \dot{p} = (pH - Hp). \quad (14)$$

The equations (12) hold good at one instant. But it can be shown from (14) that they are valid for all values of t . Consider, for example, (12a) :

$$\begin{aligned} \frac{d}{dt}(qp - pq) &= \dot{q}p - p\dot{q} + q\dot{p} - \dot{p}q \\ &= \frac{2\pi}{i\hbar} \left[(qH - Hq)p - p(qH - Hq) \right. \\ &\quad \left. + q(pH - Hp) - (pH - Hp)q \right] \\ &= \frac{2\pi}{i\hbar} \left[H(pq - qp) + (qp - pq)H \right] \quad (15) \\ &= H - H = 0, \quad (16) \end{aligned}$$

when use is made of (12a) in (15). Hence the quantum conditions (12) remain invariant for all time. So the second postulate contained in (14) does not conflict with the first as given by (12).

The p - q relations are further supplemented by Dirac by the relation²

$$-tE + Et = \frac{i\hbar}{2\pi} \quad (17)$$

There are two reasons for treating t on the same footing as q . The first is the principle of relativity which dispenses with the concept of absolute time. Hence the relative time as it affects the measurements must be bound up with a canonical conjugate as q and p are. Secondly, in the description of a particle the space time and the energy-momentum are so correlated that the energy and

time must mutually obey the same relation that the momentum and position do. This is made particularly obvious by the wave function,

$$\psi = e^{2\pi i(\tau x - \nu t)},$$

taken with the quality relations,

$$E = h\nu, \quad \hat{p} = h\hat{\tau}.$$

We have already shown that the uncertainty relation,

$$\Delta p \Delta x \sim h,$$

follows from the resolution of a function of x into wave terms. Similarly the Fourier expansion of a time function gives

$$\Delta E \Delta t \sim h.$$

It stands to reason, therefore, that a quantum condition such as (17) should be valid. Following Dirac's abstract formulation it will be shown in the next chapter that the non-commutability relations lead to the uncertainty conditions. At this stage it is enough to point out that the Hamiltonian equations follow from the Pfaffian,

$$\sum p_r dq_r - H dt. \quad (18)$$

If $H = E$, t and E can be taken as conjugate variables. $H-E$ plays the role of the Hamiltonian and the equations of motion are

$$\dot{x} = [x, H-E] \quad (19)$$

where by $[x, y]$ is meant the generalized Poisson bracket,

$$\sum_r \left(\frac{\partial (x, y)}{\partial (q_r, p_r)} - \frac{\partial x}{\partial t} \frac{\partial y}{\partial E} + \frac{\partial x}{\partial E} \frac{\partial y}{\partial t} \right). \quad (20)$$

This gives the additional quantum relations

$$\begin{aligned} tq - qt &= 0, & tp - pt &= 0, \\ Eq - qE &= 0, & Ep - pE &= 0, \\ -tE + Et &= i\hbar/2\pi. \end{aligned} \quad (21)$$

The equations (13), (14), (19), and (21) are all mutually compatible.

The diagonal matrices play a very important part in this as well as the more developed theory of Dirac's. In Heisenberg's theory the frequency term does not appear in any of the diagonal terms as is quite clear from (4). Hence for a diagonal matrix, M , the time-derivative must vanish. We know that in a Fourier series the constant term gives the time-average of the quantity represented by it. Similarly the diagonal terms of a general matrix, give the time-averages of the various stationary states. If a matrix M is such that

$$2\pi i\nu(n, m) M(n, m) = 0 \quad (22)$$

for every pair of sets, n and m and if the system is not degenerate

$$\nu(n, m) \neq 0, \text{ and } M(n, m) = 0; \quad (23)$$

and the matrix is diagonal. In case the system is degenerate so that $E_n = E_m$, it may happen that

$$v(n, m) = 0 \text{ and } M(n, m) \neq 0 \quad (24)$$

One of the major difficulties which checked the progress of the matrix theory concerned the representation of derivatives. The fundamental relations between p , q , E , t give the time-derivatives of only the canonical variables. The theory must also give us matrix representations of perfectly general functions of the variables and their derivatives. The difficulty was removed by the non-commutability relations. Consider,

$$qp - pq = i\hbar/2\pi.$$

$$\text{Hence } qp^2 - pqp = \frac{i\hbar p}{2\pi} \text{ and } pq^2 - p^2q = \frac{i\hbar p}{2\pi}.$$

$$\text{Therefore, } qp^3 - p^3q = \frac{i\hbar}{2\pi} 2p.$$

Similarly, it can be proved by successive multiplications and additions that

$$qp^n - p^nq = \frac{i\hbar}{2\pi} np^{n-1}. \quad (25)$$

Thus if $f(p)$ is any function which can be expanded in a power series of p ,

$$qf(p) - f(p)q = \frac{i\hbar}{2\pi} f'(p). \quad (26)$$

In the same manner it can be shown that

$$pf(q) - f(q)p = -\frac{i\hbar}{2\pi} f'(q). \quad (27)$$

The results (26) and (27) can be proved to be true for more general functions of p and q .

Dirac has generalized Heisenberg's postulates by stating for any two quantum magnitudes the relation,

$$\frac{i\hbar}{2\pi} [A, B] = (AB - BA) \quad (28)$$

In classical mechanics if ξ is any function of p 's and q 's,

$$\dot{\xi} = [\xi, H].$$

$$\text{Hence, } \frac{i\hbar}{2\pi} \dot{\xi} = (\xi H - H \xi) \quad (30)$$

and it follows immediately that if H is a function like ξ , it must be constant and therefore representable by a diagonal matrix when there is no degeneracy. It can be shown alternatively that if H is a diagonal matrix the relation (30) holds good.

Returning to Heisenberg's method the solution of a problem requires that Hermitian matrices p_k, q_k , be determined satisfying the exchange relations (12) and making H a diagonal matrix. If the factors

$$e^{2\pi i t (H_n - H_m)/h}$$

are associated with the matrices p_k, q_k the equations (14) are also satisfied.

If p_k and q_k do not make H diagonal one subjects the matrices to, what Dirac calls, a contact transformation. If S is any matrix, the new variables p'_k, q'_k as given by

$$p'_k = S^{-1} p_k S, \quad q'_k = S^{-1} q_k S \quad (31)$$

also satisfy the exchange relations. The property of invariance of a Poisson bracket under a contact transformation suggests by analogy the same name here for the transformation (31). The effect on the Hamiltonian is shown by

$$H' = S^{-1} H S. \quad (32)$$

When S is so chosen that H' is a diagonal matrix the problem is solved. One limitation on the choice of S is imposed by the condition that p_k, q_k as defined above, must be Hermitian. That requires S to be unitary. Hence

$$\sum_k S(n, k) S^*(m, k) = 0, \quad (33)$$

$$\sum_k S(n, k) S^*(n, k) = 1.$$

The equation (32) is the same as

$$HS = SH'.$$

The matrix terms (n, m) of the two sides follow the equality :

$$\sum_k H_{nk} S_{km} = \sum_k S_{nk} H'_{km} = S_{nm} H'_{mm} \quad (34)$$

as H' is a diagonal matrix. Corresponding to any particular value of m , by varying n , we get a set of homogeneous linear equations which have a non-trivial single solution if the discriminant vanishes, that is, if

$$| H_{rs} - \delta_{rs} H'_{mm} | = 0, \quad (35)$$

where δ_{rs} stands for the (r, s) term of the unit matrix. All the diagonal components of H satisfy that equation. For each value of m we get one column of the matrix S from (34). In the case of a degenerate system the roots H'_{mm} of the determinantal equation are not all unequal.

The perturbation theory for matrices is extremely simple in form. Let the Hamiltonian consist of the two parts

$$H = H_0 + \Delta H \quad (36)$$

of which the second part can be expanded in powers of a small parameter λ :

$$\Delta H = \lambda H_1 + \lambda^2 H_2 + \dots \quad (37)$$

Suppose that the problem is solved for H_0 which means that the coordinates and momenta make H_0 diagonal. For the general problem we have to find a unitary transformation function S so that

$$HS = SH' \quad (38)$$

H' being a diagonal matrix. We will assume the following expansion for S and H' :

$$S = S_0 + \lambda S_1 + \lambda^2 S_2 + \dots \quad (39)$$

$$H' = H_0 + \lambda W_1 + \lambda^2 W_2 + \dots \quad (40)$$

Substituting in (38), the different powers of λ give

$$S_0 H_0 = H_0 S_0 \quad (41)$$

$$H_0 S_1 + S_1 H_0 - S_1 H_0 = S_0 W_1 \quad (42)$$

etc. It is enough to discuss here the two equations (41) and (42). The first of the two is equivalent to

$$\sum_k \left\{ S_0(n, k) H_0(k, m) - H_0(n, k) S_0(k, m) \right\} = 0,$$

$$\text{or} \quad S_0(n, m) H_0(m, m) - H_0(n, n) S_0(n, m) = 0.$$

Hence if there is no degeneracy, that is, if $H_0(m, m) \neq H_0(n, n)$

$$\text{we get} \quad S_0(n, m) = 0 \quad n \neq m$$

and we can take S_0 as the unit matrix. The equation (42) now becomes

$$H_0 S_1 - S_1 H_0 + H_1 = W_1;$$

it gives the following relation between the matrix elements :

$$H_0(n, n) S_1(n, m) - S_1(n, m) H_0(m, m) + H_1(n, m) = W(n, m),$$

or, using Bohr's frequency condition,

$$h\nu(n, m) S_1(n, m) = W_1(n, n) \delta_{nm} - H_1(n, m). \quad (43)$$

It is assumed that every matrix in the expansion of H' is diagonal. From (43) we get

$$H_1(n, n) = W_1(n, n). \quad (44)$$

When $n \neq m$ the non-diagonal elements of S_1 are given by (43). But the equation fails to give the diagonal element of S_1 . Since S is to satisfy the conditions (33) we can take the diagonal elements of S_1 to be zero to the first order of approximation. The non-diagonal elements are given by

$$S_1(n, m) = \frac{H_1(n, m) (1 - \delta_{nm})}{h\nu(n, m)}. \quad (45)$$

The relation (44) is the analogue of the classical one, viz., that the time-average of the perturbation function is the perturbation energy. In Heisen-

berg's theory, the diagonal elements are time-averages as we have pointed out earlier.

Suppose however that the undisturbed state is degenerate. Let

$$H_0(n_1, n_1) = H_0(n_2, n_2) = H_0(n_3, n_3). \quad (46)$$

We cannot therefore conclude that $S_0(n, m) = 0$ when n, m are any two of the integers n_1, n_2, n_3 . We can however assume, if there is no other source of degeneracy, that barring these three rows and columns S_0 is a diagonal matrix. Moreover the condition for S_0 to be unitary makes all the elements of the diagonalized parts of the matrix equal to unity. The equation (42) provides the necessary information regarding the non-vanishing elements of S_0 and the average values of W_1 . If $n \neq n_1, n_2, n_3$ the equation (44) still holds good. For these three states we have the following homogeneous equations :

$$\left. \begin{aligned} S_0(n_1, n_1)[H_1(n_1, n_1) - W_1(n_1, n_1)] + S_0(n_1, n_2)H_1(n_2, n_1) \\ \quad + S_0(n_1, n_3)H_1(n_3, n_1) = 0 \\ S_0(n_1, n_1)H_1(n_1, n_2) + S_0(n_1, n_2)[H_1(n_2, n_2) - W_1(n_1, n_1)] \\ \quad + S_0(n_1, n_3)H_1(n_3, n_2) = 0 \\ S_0(n_1, n_1)H_1(n_1, n_3) + S_0(n_1, n_2)H_1(n_2, n_3) \\ \quad + S_0(n_1, n_3)[H_1(n_3, n_3) - W_1(n_1, n_1)] = 0 \end{aligned} \right\} \quad (47)$$

Eliminating $S_0(n_1, n_1)$, $S_0(n_1, n_2)$ and $S_0(n_1, n_3)$ we get a determinantal equation for the value of $W_1(n_1, n_1)$. The same equation gives $W_1(n_2, n_2)$ and $W_1(n_3, n_3)$. Hence the three missing terms of W_1 come as the roots of a cubic.

2. THE HARMONIC OSCILLATOR

We will take up in this the simple problem of the harmonic oscillator^a and show how it has been worked out in the matrix theory. The solution of this problem in accordance with Schrödinger's theory has been already introduced earlier. The matrix components can also be evaluated from Schrödinger's ψ -functions as was first pointed out by him. The general connections between the wave functions and the matrices will be shown in the next section.

For an oscillator,

$$H = \frac{1}{2m_0} p^2 + \frac{1}{2} \mu q^2, \quad p = m\dot{q}, \quad (48)$$

$$\text{and} \quad \ddot{q} + (2\pi\nu_0)^2 q = 0, \quad (49)$$

$$\text{where} \quad (2\pi\nu_0)^2 = \mu/m_0. \quad (50)$$

Assuming the matrix

$$q(n, m) e^{2\pi i v(n, m) t}$$

for q we get on substitution in (49), when it is treated as a matrix equation*

$$\{v^2(n, m) - v_0^2\} q(n, m) = 0. \quad (51)$$

$$\text{Hence } q(n, m) = 0, \text{ unless } v(n, m) = \pm v_0. \quad (51')$$

The exchange relation

$$qp - pq = \frac{ih}{2\pi} \quad (52)$$

becomes for the diagonal terms

$$\begin{aligned} \frac{ih}{2\pi} = \sum_k \left\{ q(nk) \cdot m_0 \cdot 2\pi i v(k, n) q(kn) \right. \\ \left. - m_0 2\pi i v(n, k) q(nk) q(kn) \right\}. \end{aligned} \quad (53)$$

Since

$$v(n, k) = -v(k, n) \text{ and } q(kn) = q^*(n, k) \quad (54)$$

$$\sum_k v(n, k) |q(kn)|^2 = -\frac{h}{8\pi^2 m_0} \quad (55)$$

We can now evaluate the energy of the oscillator in a stationary state from (48).

$$H(nn) = 2\pi^2 m_0 \sum_k \{v_0^2 + v^2(n, k)\} |q(nk)|^2. \quad (56)$$

The three important equations on which the subsequent discussion hangs are (51), (55) and (56). From (51'), $q(n, m) = 0$ unless

$$H(nn) - H(mm) = \pm h v_0. \quad (57)$$

If (55) is to hold good it is necessary that all $q(n, m)$ should not be zero. Hence for any n there exists an m such that (57) is satisfied. In the absence of degeneracy there are only two values of m satisfying the last condition. Let us take

$$\begin{aligned} H(nn) - H(n-1, n-1) &= h v_0, \\ H(nn) - H(n+1, n+1) &= -h v_0. \end{aligned}$$

There is no loss of generality in this as the order in which terms are arranged in the matrix form has not been specified. We have now

$$\begin{aligned} v(n, n-1) &= v_0, v(n, n+1) = -v_0, \\ q(n, n-1) &\neq 0, q(n, n+1) \neq 0, \end{aligned}$$

* Henceforth we will find it convenient to denote the matrix elements $q(n, m)$ etc. as $q(nm)$ etc.

all other values of v and q -components vanishing. Hence, from (56)

$$H(n, n) = 4\pi^2 m_0 \nu_0^2 \{ |q(n, n-1)|^2 + |q(n, n+1)|^2 \}. \quad (58)$$

The equation (45) now reduces to

$$\frac{-h}{8\pi^2 m_0} = \pm \nu_0 \{ |q(n, n-1)|^2 - |q(n, n+1)|^2 \}. \quad (59)$$

From (58) and (59) it can be shown by considering a series of difference equations, that the equations

$$H(n_k, n_k) = h\nu_0 (k + \frac{1}{2}), \quad (60)$$

$$|q(n_k, n_{k+1})|^2 - |q(n_k, n_{k-1})|^2 = \frac{h}{8\pi^2 m_0 \nu_0}, \quad (61)$$

$$|q(n_k, n_{k+1})|^2 = \frac{h(k+1)}{8\pi^2 m_0 \nu_0} \quad (62)$$

hold good. The ordering of the suffixes is done so that n_0 corresponds to the ground state; n_1, n_2, \dots, n_k are the successive higher states of energy. The q -matrix is of the following form if we take n_0, n_1, n_2, \dots in the natural order of integers.

$$\begin{vmatrix} 0 & , & q(01) e^{2\pi i \nu(0,1)t} & , & 0 & , & 0 & , & \dots \\ q(10) e^{2\pi i \nu(1,0)t} & , & 0 & & q(12) e^{2\pi i \nu(1,2)t} & , & 0 & , & \dots \\ 0 & , & q(21) e^{2\pi i \nu(2,1)t} & , & 0 & , & q(23) e^{2\pi i \nu(2,3)t} & , & 0 & , & \dots \\ \dots & & \dots & & \dots & & \dots & & \dots \\ \dots & & \dots & & \dots & & \dots & & \dots \end{vmatrix} \quad (63)$$

The final results may now be expressed in the familiar form :

$$H(n, n) = h\nu_0 (n + \frac{1}{2}), \quad (60')$$

$$|q(n, n+1)|^2 = \frac{h(n+1)}{8\pi^2 m_0 \nu_0}, \quad (62')$$

$$\nu(n+1, n) = \nu_0. \quad (64)$$

The theory of perturbation is illustrated by the anharmonic oscillator for which we take

$$H = H_0 + \lambda q^3$$

$$\text{where } H_0 = \frac{1}{2} \frac{p^2}{m} + \frac{1}{2} \lambda q^2, \quad \lambda q^3 \equiv H_1.$$

The matrix of H_1 is given by

$$H_1(n, m) = \sum_k \sum_l \lambda [q(n, k) q(k, l) q(l, m)]$$

To carry out the summation for k and l we have to note that for any n , the non-zero members of the q -matrix are only $q(n-1, n)$ and $q(n, n+1)$.

Hence we need take into consideration merely the following sets of values for k, l, m .

$$\begin{aligned} k: & (n-1), & (n+1); \\ l: & (n-2), & n, & (n+2); \\ m: & (n-3), & (n-1), & (n+1), & (n+3). \end{aligned}$$

It is obvious that the non-zero terms of H_1 are of the form

$$H_1(n, n-3), H_1(n, n-1), H_1(n, n+1), H_1(n, n+3).$$

The diagonal terms of H_1 are all zero and, therefore, the first order perturbation energy is zero. The calculations have to be pushed to the second order of approximation for getting the energy arising out of H_1 .

3. THE CONNECTION BETWEEN WAVES AND MATRICES

Schrödinger and Eckart have independently shown that the same physical facts are represented by Heisenberg's matrices and the Schrödinger wave functions. The equivalence follows from the Fourier expansion theorem by means of a set of normalized orthogonal functions and the curious representation of a momentum co-ordinate by a differential operator.

Let u be any classical function. Then we can associate a matrix with it through a set of normalized orthogonal functions ϕ_k . For consider the function $u \phi_n$. Let us write the expansion of this function as

$$u \phi_n = \sum_k u(k, n) \phi_k, \quad u \phi_n^\times = \sum_k u(n, k) \phi_k^\times. \quad (65)$$

The values of the co-efficients $u(n, k)$ satisfy the condition,

$$u(k, n) = \int u \phi_n \phi_k^\times d\tau. \quad (66)$$

Similar matrix elements can be associated with any other function v . These arrays of functions satisfy the matrix rule of multiplication for

$$\begin{aligned} u v(n, m) &= \int u v \phi_n^\times \phi_m d\tau \\ &= \int v \sum_k u(n, k) \phi_k^\times \phi_m d\tau \\ &= \sum_k \sum_j u(n, k) v(j, m) \int \phi_k^\times \phi_j d\tau \\ &= \sum_k u(n, k) v(k, m). \end{aligned} \quad (67)$$

That Heisenberg's matrices follow from the solutions of Schrödinger's equations has been shown by considering the integral,

$$I_{nm} = \int \phi_n^\times f(q, -\frac{i\hbar}{2\pi} \partial/\partial q) \phi_m d\tau. \quad (68)$$

This integral is interpreted as the (n, m) element of the matrix for $f(q, p)$. ϕ_n^* and ϕ_m are now solutions of Schrödinger's equation of the system. The volume integration is carried over all space. The Heisenberg matrices satisfy the condition,

$$pq - qp = -\frac{ih}{2\pi}.$$

Using (68) we have

$$\begin{aligned} [pq - qp]_{nm} &= \int \phi_n^* \left(-\frac{ih}{2\pi} \frac{\partial}{\partial q} q + q \frac{ih}{2\pi} \frac{\partial}{\partial q} \right) \phi_m d\tau \\ &= -\int \frac{ih}{2\pi} \phi_n^* \phi_m d\tau = -\frac{ih}{2\pi} \delta_{nm}. \end{aligned} \quad (69)$$

Hence the matrices given by the rule obey the exchange relations. We can show that they satisfy the matrix equations of motion, viz.,

$$-\frac{ih}{2\pi} \dot{q} = Hq - qH, \quad -\frac{ih}{2\pi} \dot{p} = Hp - pH. \quad (70)$$

For, consider the function $Hq\phi_m$:

$$\begin{aligned} Hq\phi_m &= H \sum q(km) \phi_k = \sum q(km) H\phi_k \\ &= \sum q(km) E_k \phi_k, \end{aligned}$$

where E_k is the characteristic energy for H_k . Similarly,

$$qH\phi_m = qE_m\phi_m = E_m q\phi_m.$$

Now the value of the integral,

$$\begin{aligned} &\int \phi_n^* (Hq - qH) \phi_m d\tau \\ &= \int \phi_n^* \sum q(km) E_k \phi_k d\tau - E_m \int \phi_n^* q \phi_m d\tau. \end{aligned}$$

The numbers $q(km)$ are constant; so also E_k . Therefore

$$\begin{aligned} \int \phi_n^* (Hq - qH) \phi_m d\tau &= (E_n - E_m) q_{nm} \\ &= h\nu_{nm} q_{nm} = -\frac{ih}{2\pi} \dot{q}_{nm}. \end{aligned} \quad (71)$$

The matrices defined by Schrödinger's functions satisfy the equation

$$Hq - qH = -\frac{ih}{2\pi} \dot{q}.$$

Similarly it can be shown that the other equation of (70) is also consistent with Schrödinger's definition of matrices.

The weaknesses and limitations of the wave theory have been already indicated. The correspondence between that theory and Heisenberg's is so close, that none of its defects could be removed by working out problems in

terms of matrices. One unsatisfactory feature of the wave treatment was the postulated equivalence of a momentum coordinate and a differential operator. This was originally suggested by an analogy from the mathematical theory of optics. It is interesting that the equivalence is due to a connection between the Poisson brackets and the exchange relations. Dirac began with a generalization of these relations as given by (28). It introduced what he first called *q*-numbers and *c*-numbers, the former satisfying the non-commutability relations and the latter being subject only to the properties of complex numbers and functions. That has very much enlarged the scope of the theory. It has presented more generalized wave functions and the new perturbation theory is made applicable not only to the energy function but to other dynamical functions as well.

4. PROBABILITY POSTULATES

Certain fundamental considerations regarding the part played by probability in quantum mechanics must be presented here. In the classical theory of probability some writers introduce the concept by considering the ratio of the equally probable favourable possibilities of an event to all the equally probable possibilities. Those who are dissatisfied with this procedure on account of the phrase 'equally probable', which implies an *a priori* knowledge, give the other definition which depends upon frequencies and statistical distributions. In wave mechanics, when we talk of probabilities we mean statistical frequencies and not *a priori* probabilities. For observations interpret the probability in statistical terms only. Wherever the concept is used theoretical predictions or inferences are made exclusively for large numbers. In certain processes if we still continue to think of individual particles that is confined to questions regarding the energy and momentum. As the matrices and waves are connected by the probability interpretation and as the calculus of probabilities in the quantum theories of matter and radiation is subject to laws different from the classical we give below a set of postulates centred round that concept. A clear statement of them has been provided by Ruark³ and Urey.

"(I) There always exists a function $\phi(x, y, F_1, F_2)$ called the probability amplitude such that

$$\phi(xy, F_1, F_2) \phi^*(x, y, F_1, F_2) = I^j(xy, F_1, F_2)." \quad (72)$$

Let us denote a fixed value of a certain dynamical variable F_2 by y_0 and enquire: what is the probability that another variable F_1 has a value lying between x_0 and $x_0 + \delta x_0$? The answer is provided by $P(x_0, y_0, F_1, F_2) \delta x_0$. Similarly, if x_0 is the fixed value of F_1 the probability that F_2 lies between y_0 and $y_0 + \delta y_0$ is $P(x_0, y_0, F_1, F_2) \delta y_0$.

"(II) The function $P(x, y, F_1, F_2)$ must satisfy the conditions.

$$\begin{aligned} P(x_0, y_0, F_1, F_2) &= 1, & \text{if } x_0 &= y_0, \\ &= 0, & \text{if } x_0 &\neq y_0. \end{aligned} \quad (73)$$

For if it is definitely known that F_1 has a fixed value x_0 the probability that it has at the same time any other value must be zero. This is certainly obvious. Before performing a certain experiment suppose that the probability $F_1 = x$ is P . Once it is known that $F_1 = x_0$ at a certain instant, if our finding is to have any value it is necessary that, on performing the experiment again, all relevant disturbances being properly controlled, the probability that $F_1 = x$ must be a certainty. Hence the effect of a measurement is to disturb the probability function. This fact has to be taken account of in framing the postulates of probabilities. The function as given by (73) is of the same type as the Kronecker delta which is generalized in quantum mechanics, for continuous representations into the function $\delta(x)$:

$$\begin{aligned}\delta(x) &= 0, & x &\neq 0 \\ \int_{x_1}^{x_2} \delta(x) dx &= 1,\end{aligned}\tag{74}$$

where x_1 and x_2 are any numbers so that $x_2 > 0 > x_1$.

(III) According to the classical law of combination of probabilities

$$P(x, z, F_1 F_3) dz = \int P(x, y, F_1 F_2) dy P(y, z, F_2 F_3) dz.\tag{75}$$

Given the value x of F_1 we want to find out the probability that F_3 lies between z and $z + \delta z$ it being assumed that F_2 may have any value y . The answer provided by (75) is not true in quantum mechanics. The equation (75) is replaced by

$$\phi(x, z, F_1 F_3) = \int \phi(x, y, F_1 F_2) \phi(y, z, F_2 F_3) dy,\tag{76}$$

and the required probability is obtained from (72).

(IV) It is assumed that the probability functions are dependent exclusively on the functional forms $F_1(p, q)$ and $F_2(p, q)$, being independent of charges and other special properties of the mechanical systems.

It is on account of the form (76) that the wave functions can be, in many problems, complex as functions of x, y, z .

Suppose¹ that the probability of transition between two states $n \rightarrow l$ is indicated by $|S'_{nl}|^2$. Let there be a further disturbance causing transition of the type $l \rightarrow m$, the probability being $|S''_{lm}|^2$. Hence the probability that a system which was in the state n initially might be in the state m finally is

$$|\sum_l S'_{nl} S''_{lm}|^2.\tag{77}$$

This probability must be distinguished from the probability

$$|\sum_l S'_{nl}|^2 |S''_{lm}|^2,\tag{78}$$

which would be true on the classical theory. The question now arises : what is the difference in physical significance between (77) and (78)? If we take (77) this probability is based on the understanding that no measurements are made in the intermediate l state. On the other hand (78) shows the effect of an actual measurement made to find out how the transition takes place from n to m . Unless the intermediate state is observed the term $|S''_{lm}|^2$ has no place in our theoretical estimate of the transition $n \rightarrow m$.

REFERENCES

- (1) HEISENBERG, W. "*The Physical Principles of the Quantum Theory*", pp. 59-62, 107-41 (1930).
- (2) BIRTWISTLE, G. "*The New Quantum Mechanics*," pp. 70-99 (1928).
- (3) RUARK and UREY, "*Atoms, Molecules and Quanta*," pp. 594-616 (1930).

VI STATES AND OBSERVABLES

1. A Mathematical Analogy. 2. The Calculus of Operators. 3. The Postulates and Principles of Dirac's Theory. 4. The Uncertainty Principle. 5. The Electron Spin.

1. A MATHEMATICAL ANALOGY

The quantum mechanics as expounded by Dirac is a theory of states and observables founded on the concept of probability. By way of introducing the new philosophy a mathematical analogy is presented here. When the purpose is served the analogy must not be stretched further. Moreover once a good acquaintance is cultivated with the basic terms and processes of the theory it is no longer necessary to consider any analogies since the formal mathematical structure is itself extremely simple and beautiful.

Suppose we have to describe an ellipse. In Cartesian co-ordinates it may be given as

$$ax^2 + 2hxy + by^2 = 1, \quad ab - h^2 > 0, \quad (1)$$

taking the origin at the centre. In polar co-ordinates every point on the ellipse becomes known when the value of the radius vector r is given for every θ . There exists a different unit vector for each value of θ if $0 \leq \theta \leq 2\pi$. Let us call each of the vectors a state. Corresponding to every direction in the plane of the ellipse there is a state; and corresponding to each state there is a magnitude r which we call the value of the observable (ellipse) for that state. When we know the values of the observable for all possible states the ellipse becomes completely determined.

Fortunately all this is not needed for a complete knowledge of the ellipse. It is enough to know the directions of the principal axes and corresponding values of the observable. The ellipse can be constructed with this much information only. The four directions of the principal axes give four eigenstates and the corresponding values of r give the eigenvalues. Sufficient information of the ellipse is provided by the eigenstates and the eigenvalues.

By a transformation of axes,

$$\begin{aligned} x &= X \cos \theta - Y \sin \theta \\ y &= X \sin \theta + Y \cos \theta, \end{aligned} \quad (2)$$

we have from (1)

$$\frac{X^2}{A^2} + \frac{Y^2}{B^2} = 1 \quad (1')$$

where $A^{-2} = a \cos^2 \theta + 2h \sin \theta \cos \theta + b \sin^2 \theta,$ (3a)

$$B^{-2} = a \sin^2 \theta - 2h \sin \theta \cos \theta + b \cos^2 \theta, \quad (3b)$$

$$0 = (b - a) \sin \theta \cos \theta + h \cos^2 \theta. \quad (3c)$$

The form (1') gives directly the eigenstates which are four unit vectors parallel to Ox and Oy and the eigenvalues A and B .

Between (1) and the matrix

$$\begin{pmatrix} a & h \\ h & b \end{pmatrix} \quad (4)$$

a one-to-one correspondence can be established at once. If the equation (1) represents a class of ellipses distinguished by different values of a, h, b the matrix (4) can also be treated as an efficient representative of the same class of ellipses. The operation of transformation (2) gives the new matrix

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (5)$$

and this is equal to

$$\begin{pmatrix} a \cos^2 \theta + 2h \sin \theta \cos \theta + b \sin^2 \theta & (b - a) \sin \theta \cos \theta + h \cos 2\theta \\ (b - a) \sin \theta \cos \theta + h \cos 2\theta & a \sin^2 \theta - 2h \sin \theta \cos \theta + b \cos^2 \theta \end{pmatrix} \quad (6)$$

Hence when the conditions (3) are fulfilled, we have the diagonal matrix

$$\begin{pmatrix} A^{-2} & 0 \\ 0 & B^{-2} \end{pmatrix} \quad (4')$$

as a representative of the observable. If we denote the matrix (4) by α and the matrices

$$\begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}, \quad \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

by S^{-1} and S it will be noted that

$$S^{-1} S = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix};$$

also the matrix (6) is equivalent to $S^{-1} \alpha S$. The condition (3c) diagonalizes the matrix giving the form (4') corresponding to (1'). The eigenstates and eigenvalues are given by the set of equations

$$\begin{aligned} (a - \lambda)x + hy &= 0, \\ hx + (b - \lambda)y &= 0, \end{aligned} \quad (7)$$

which express the conditions for r to be a maximum or a minimum. Eliminating x and y we get the determinantal equation

$$\begin{vmatrix} a - \lambda & h \\ h & b - \lambda \end{vmatrix} = 0 \quad (8)$$

of which the roots are $\lambda = A^{-2}, B^{-2}$. When these values of λ are substituted in (7) the eigenstates follow from the ratio y/x .

In the above we have shown how an observable such as an ellipse can be represented by a matrix. The significance of diagonalizing the matrix is pointed out in the notation of transformation of axes. The meanings of eigenvalues and eigenstates are made explicitly clear. The latter are represented by mutually perpendicular vectors. When two or more eigenstates correspond to the same eigenvalue there is said to be a degeneracy. A as well as B corresponds to two different eigenstates and thus the concept of degeneracy is illustrated in the case of the ellipse.

Suppose we have to consider another observable, also an ellipse. Call it α' and let the matrix

$$\begin{pmatrix} a' & h' \\ h' & b' \end{pmatrix} \quad (9)$$

stand for it. The states of α can be represented in terms of two eigenstates, ψ_A and ψ_B of α by an equation,

$$\psi(\alpha) = \psi_A \cos \theta + \psi_B \sin \theta \quad (10)$$

which is the usual equation for expressing one vector in terms of two perpendicular vectors in the same plane. Now the most important question is whether by a suitable transformation of axes both the observables can be put in the form (1') or (4'). Obviously this is possible only if the two ellipses have the same principal axes. When the two observables are simultaneously in the diagonal form

$$\begin{pmatrix} A^{-2} & 0 \\ 0 & B^{-2} \end{pmatrix} \begin{pmatrix} A'^{-2} & 0 \\ 0 & B'^{-2} \end{pmatrix} = \begin{pmatrix} A'^{-2} & 0 \\ 0 & B'^{-2} \end{pmatrix} \begin{pmatrix} A^{-2} & 0 \\ 0 & B^{-2} \end{pmatrix}$$

$$\text{or} \quad S^{-1} \alpha S \cdot S^{-1} \alpha' S = S^{-1} \alpha' S^{-1} \alpha S$$

$$\text{or} \quad S^{-1} \alpha \alpha' S = S^{-1} \alpha' \alpha S.$$

$$\text{Hence} \quad \alpha \alpha' = \alpha' \alpha. \quad (11)$$

Two observables whose matrices commute are called compatible. Their analogues are ellipses whose principal axes are common. It is easy to trace the steps backward and prove that when the condition (11) does not hold good the two ellipses cannot be simultaneously expressed in the form (1'). For non-compatible observables α and α' , (11) is not true but they can be shown to be functions of some fundamental observables for which Dirac has postulated relations of the form

$$pq - qp = -i\hbar/2\pi.$$

Actually in the quantum theory the space-time co-ordinates themselves along with the other dynamical variables are the observables and the space in

which the observables are represented is a Hermitian space of an infinite number of dimensions. The analytical operations of this complex space which give Dirac's theory its distinctive features have very few analogues among those of the Euclidean space. It is not advantageous therefore to pursue the analogy further.

2. THE CALCULUS OF OPERATORS

The mathematical theory underlying the quantum mechanical generalizations by Dirac is very often described as a calculus of operators. An operator is a rule for replacing a function by a function. In Dirac's theory the functions on which the operators act are wave functions, ψ and ψ^* . They do not represent waves but as they are of the same class as Schrödinger's function this name is used for them. The operators having a physical significance are the observables which are so called because the numerical values that they represent can be identified with the results of measurement.

Let us denote an operator by T . Suppose

$$Tf(x) = g(x),$$

that is, the result of the operation is to convert $f(x)$ into $g(x)$. In particular, let $g(x)$ be the derivative of $f(x)$. If S is any other operator which stands for multiplication by x ,

$$STf(x) = Sf'(x) = xf'(x).$$

Obviously, the operation ST is not the same as TS . For

$$TSf(x) = Tx f(x) = f(x) + x f'(x).$$

Hence

$$(TS - ST)f(x) = f(x),$$

or

$$TS - ST = 1.$$

This introduces the simplest of all operators, the identity. Consider now an operator A which transforms $f(x)$ into $g(y)$

$$Af(x) = g(y). \quad (12)$$

The equation may be of the form

$$g(y) = \int A(x, y) f(x) dx \quad (13)$$

or if x and y each take n discrete values

$$g(y) = \sum_{x=1}^n a(yx) f(x). \quad (13')$$

It is well known that the transformations of this kind form a group. [The reader is introduced to the abstract idea of a group in the discussion on force.] So if A and B are two operators acting in a specific order on any function the result is the same as that of a certain operator of the same group. If

$$\begin{aligned}
 h(x) &= \sum b(zy) g(y), \\
 \text{then } h(z) &= \sum \sum b(zy) a(yx) f(x) \\
 &= \sum c(zx) f(x).
 \end{aligned}$$

Hence the matrix law of multiplication is satisfied by A, B, C ; for

$$c(zx) = \sum_y b(zy) a(yx). \quad (14)$$

For a full working knowledge of any operator A it is enough to find out what vector $f(x)$ is left by it unchanged in direction. The linear transformation associated with A gives us a matrix. The characteristic vectors $f(x)$ which remain invariant are given by

$$Af = af \quad (15)$$

$$\text{or } \sum a(yx) f(x) = a f(y). \quad (16)$$

Since x and y take only integral values from 1 to n , the condition that (16) may have a non-trivial solution is, in the usual notation,

$$|a_{rs} - a\delta_{rs}| = 0. \quad (17)$$

The roots are the eigenvalues of A ; for each root we get a vector $f(x)$ from the homogeneous equations. When all the roots are unequal there are n orthogonal vectors defined by the roots. If two or more roots are equal we can still construct a set of n mutually orthogonal vectors. But the eigenvalues associated with some of the directions are equal.

The vectors $f(x)$ and $g(y)$ represent physical realities independent of the co-ordinate system chosen. Suppose therefore that in a new system the vectors are represented by $F(x)$ and $G(y)$. The law of transformation which was originally

$$\begin{aligned}
 Af(x) &= g(y) \\
 \text{now becomes } A'F(x) &= G(y).
 \end{aligned} \quad (17)$$

What is the connection between A and A' ? To answer this question suppose that the two co-ordinate systems are connected by a transformation T so that

$$\begin{aligned}
 Tf(x) &= F(x) \\
 \text{or } f(x) &= T^{-1}F(x);
 \end{aligned} \quad (18)$$

$$\text{and } Tg(y) = G(y).$$

$$\text{Therefore } AT^{-1}F(x) = g(y)$$

$$\text{or } TAT^{-1}F(x) = G(y).$$

$$\text{Hence } A' = TAT^{-1}. \quad (19)$$

Suppose that the matrix A is referred to its characteristic vectors as axes. Let the n vectors be given by $\phi(xu)$ where $u = 1, 2, \dots, n$. In the new co-ordinate system they will be described by the functions $\phi'(yu)$:

$$\phi'(yu) = \sum_x t_{yx} \phi(xu) \quad (20)$$

where t_{yu} is the matrix of the transformation. Since

$$\phi'(yu) = \delta_{yu} \quad (21)$$

it follows that $t^{-1}_{yx} = \phi(xu)$.

Hence in the new co-ordinate system A is replaced by A' :

$$a'_{yu} = \sum t_{yx} a_{xz} t^{-1}_{zu} = \sum t_{yx} a_u \phi(xu) = a_u \delta_{yu}. \quad (22)$$

The new matrix is diagonal and its non-zero members are the eigenvalues of A .

Usually the components of the vectors $f(x)$ are complex numbers. Hence the length of the vector is L so that

$$L^2 = \sum f(x) f^*(x) \quad (23)$$

The condition for the invariance of L requires the transformations to be unitary.

In Dirac's theory the dynamical state of a system is represented by a vector ϕ . But the state has no longer the same meaning as in the classical theory. We are accustomed to talk about the state of a system at an instant as given by the values of certain variables. In quantum mechanics what is meant by the state is the familiar vector ϕ which by itself has not a physical meaning. But we can construct the expressions of probability current and density from ϕ . All the mechanical possibilities that are inherent in the system are brought out by these expressions and also by the integrals

$$\int_{\tau} \phi_m^* F \phi_n d\tau$$

where F is any observable and ϕ_m and ϕ_n are the same or two different states. As the state vector ϕ has complex components the whole theory requires two separate conjugate spaces in which the conjugate imaginaries ϕ and ϕ^* are defined. These spaces are Hermitian and they may be of infinitely many dimensions. The operators simply play the part of relations of connection for two state vectors. By the normalization of a vector is meant that the length is unity. For one variable, the condition of normalization is

$$\int_{-\infty}^{\infty} \phi^*(x) \phi(x) dx = 1.$$

Suppose that a dynamical system is given by the normalized state vector $\phi(x)$, so that $\phi^*(x) \phi(x) dx$ is the probability that the system is between x and $x + dx$. If we want to know the probabilities of the various values of energy it is necessary to change the co-ordinate system so that the new principal axes are those of energy. Accordingly a transformation function $t(W, x)$ must be found so that the new state vector $\phi(W)$ is given by

$$\phi(W) = \int t(W, x) \phi(x) dx, \quad (24)$$

From the analogue of (21) for continuous x ,

$$t^{-1}(W, x) = \psi(x, W) \quad (25)$$

which gives us the vectors of energy with reference to the first co-ordinate system. Obviously $\psi(x, W)$ must be the characteristic vectors of energy as given by

$$H\psi = W\psi, \quad (26)$$

where W is a constant. (26) is in fact Schrödinger's equation. The reciprocal of the transformation which leads us from the state vector of position to that of energy is the Schrödinger wave function. Both position and energy cannot be simultaneously diagonalized on account of the observables being incompatible. As the transformation $t(W, x)$ must be unitary, (25) is equivalent to

$$t(W, x) = \psi^*(x, W) \quad (25')$$

and

$$\int \phi^*(W) \phi(W) dW = 1. \quad (27)$$

Hence the probability that the energy lies between W and $W + dW$ is $\phi^*(W) \phi(W)$. This is the answer to the question: if $\psi(x) \psi^*(x) dx$ is the probability that the system lies between x and $x + dx$ what is the probability that its energy lies between W and $W + dW$?

A similar question may be put regarding the momentum also. In this case in place of (26) we have

$$p\psi = p_0\psi \quad (28)$$

p_0 being the eigenvalue of p and ψ the state vector for momentum. In view of the exchange relations between p and x we may write (28) as

$$-\frac{i\hbar}{2\pi} \frac{\partial \psi}{\partial x} = p_0 \psi \quad (29)$$

and obtain, using the normalizing factor,

$$\psi(p_0, q) = e^{2\pi i p_0 x / \hbar} (1/\hbar)^{1/2} \quad (30)$$

The momentum vector $\phi(p_0)$ in the new co-ordinate system is

$$\begin{aligned} \phi(p_0) &= \int \psi^*(p_0, x) \psi(x) dx \\ &= \hbar^{-1/2} \int \psi(x) e^{-2\pi i p_0 x / \hbar} dx. \end{aligned} \quad (31)$$

So the probability that the momentum lies between p_0 and $p_0 + \delta p_0$ is $\phi^*(p_0) \phi(p_0) \delta p_0$ when the probability that the position lies between x and $x + \delta x$ is $\psi^*(x) \psi(x) \delta x$.

The wave functions of Dirac's theory provide the probability amplitudes as required by the first postulate of the calculus of probabilities. If we look into the remaining postulates as stated in Chap. V they will be found to be

fulfilled by the transformation functions. Any two observables which are not compatible with each other are symmetrically connected by a transformation which lead to the required probability function, symmetrical with respect to the two. Regarding the second postulate we know from (24) and (25) that

$$\int \psi^*(x, W) \psi(x, W_1) dx = \phi(W) = \delta W W_1 \quad (32)$$

if $\psi(x, W_1)$ is a solution of Schrödinger's equation (26). In the most general case can similarly be proved the relation analogous to (32). The fact that the probability amplitudes are linear transformation functions forming a group accounts for the third postulate. As the determination of these transformation functions is independent of the total charge or the form of the energy functions and such characteristics, the fourth postulate also is automatically satisfied.

To find out the probability relation between the momentum and energy we have to restate the operational equation

$$H\psi = W\psi$$

differently. It is assumed that originally the state vector for momentum is known. Hence ψ is expected to be a function of p and W . The operator H is therefore expressed as

$$H(p, q) \equiv H(p, \frac{i\hbar}{2\pi} \partial/\partial p). \quad (33)$$

This is permissible since it is consistent with the relation

$$pq - qp = -\frac{i\hbar}{2\pi}.$$

The eigenfunctions of the operator (33) are similar to the wave functions which satisfy the Schrödinger equation.

3. THE POSTULATES AND PRINCIPLES OF DIRAC'S THEORY

Having acquainted the reader with the terminology and procedure of Dirac's theory we may now state, at the risk of some repetition, the fundamental postulates on which quantum mechanics is based. The theory divides itself into two parts. In the first part one deals with the connection between the components of an atomic system at any instant and, in the second, with the progress of motion. Historically the first fundamental postulate is that concerning the equality of a commutator with the corresponding Poisson bracket of the classical theory. If $F_1(p, q)$ and $F_2(p, q)$ are two classical functions,

$$F_1 F_2 - F_2 F_1 = \frac{i\hbar}{2\pi} [F_1, F_2]. \quad (A)$$

For the first part of the theory the classical relations of the Poisson brackets provide the quantum conditions

$$pq - qp = -i\hbar/2\pi, \text{ etc.}$$

As regards the other part the Hamiltonian equations suggest quantum relations of the type

$$\frac{i\hbar}{2\pi} F_1 = F_1 H - H F_1, \text{ etc.}$$

The analogues of the classical dynamical variables are operators. They are called observables. In addition to those which originate in classical mechanics there are other observables such as the electron spin and the permutation variable which are of the quantum origin. Compatible observables commute; non-compatibles do not. The theory of states and observables is developed in order to get measurable effects for observables comparable in the limit with those of the classical theory. The representation of states is by vectors in a Hermitian space satisfying the principle of superposition. It means that the only possible connections between the states of a dynamical system are linear as given by

$$x_1 \psi_1 + x_2 \psi_2 + \dots + x_n \psi_n = 0, \quad (\text{B})$$

where x_1, x_2, \dots, x_n are numbers and $\psi_1, \psi_2, \dots, \psi_n$ are all vectors. As the operations are all linear in the theory, if ψ_1 and ψ_2 are two solutions of an operational equation so also is

$$c_1 \psi_1 + c_2 \psi_2,$$

where c_1 and c_2 are constants. Even the transformation equations of the form

$$\phi(q, t) = h^{-1/2} \int \phi(p, t) e^{2\pi i p q t / \hbar} dp$$

are examples of the principle of superposition. It is interesting to note that if the relation (B) holds good at one instant it is valid for all time. The co-efficients x_1, x_2 , etc. do not change but the states may change with time without violating the original form of superposition. Two features of this principle are noteworthy. If the same state is superposed on itself no new state results. Secondly any state that is obtained by superposition from two or more different states can be decomposed into any one or more of these states for the prediction of experimental results. This is permitted by the probability interpretation that is attached to the states. What is really novel about the superposition principle may be put this way. If ϕ is a state, intermediate between ψ_1 and ψ_2 , it is not true to say that the result of ϕ would be intermediate between those of ψ_1 and ψ_2 . For the vectors, by themselves, have no physical meaning. But it is right to say that the probability asso-

ciated with ψ lies between the probabilities for ψ_1 and ψ_2 . In the exposition as given by Dirac the following physical hypotheses enter.

"In the special case when the result of a particular observation made on the system in a particular state is with certainty one particular number, α , say, then the Hermitian operator, α say, representing the observable that is measured and the ψ —vector, ψ_α say, representing the state are connected by the equation

$$\alpha \psi_\alpha = \alpha \psi_\alpha. \quad (C)$$

Conversely, if this equation holds, a measurement of the observable represented by α made on the system in the state represented by ψ_α is certain to lead to the result α "

(Dirac : 'Quantum Mechanics' (1935), p. 30)

"If the measurement of the observable represented by α , for the system in the state represented by ψ_x is made a large number of times, the average of all the results obtained will be

$$\phi_x \alpha \phi_x \quad (D)$$

provided ϕ_x and ϕ_x are normalized."

(Dirac : 'Q.M.' (1935) p. 43)

In Dirac's statement of (D) ϕ_x stands for ψ_x^\times .

If ψ_1, ψ_2, \dots etc. are orthogonal normalized functions let

$$\psi_r^\times \psi_s = A + iB. \quad (34)$$

Then

$$\psi_s^\times \psi_r = A - iB. \quad (34')$$

This is a mathematical postulate of the theory. These relations are used to define the conjugate observable α^\times . It is given by

$$\psi_r^\times \alpha^\times \psi_s = c - iD, \quad (35)$$

if

$$\psi_s^\times \alpha \psi_r = c + iD. \quad (35')$$

In the above A, B, C, D appear for real numbers. The complex observables for which the eigenvalues are complex numbers have no physical significance. In quantum mechanics we deal with real observables only.

Let $\xi_1, \xi_2, \dots, \xi_m$ be all the independent observables of a system every two of which commute with each other. Then it is possible to put them all simultaneously in the diagonal form. It can now be shown that for any other observable α to be diagonalized in the ξ -representation it is necessary and sufficient that α be a function of ξ 's. The transformation theory connects up the ξ -state vectors with the η -state vectors where $\eta_1, \eta_2, \dots, \eta_m$ is another complete set of independent but compatible observables. The connecting equation is

$$\psi(\eta') = \sum_{\xi'} \psi(\xi') (\xi' | \eta') \quad (36)$$

or if ξ are continuous,*

$$\psi(\eta') = \int \psi(\xi') d\xi' (\xi | \eta'). \quad (36')$$

Given the state vector $\psi(\xi)$, the transformation function must be obtained by solving the equation

$$\eta\psi = \eta'\psi. \quad (37)$$

It has been already shown in (25) that this ψ is the reciprocal of $(\xi | \eta)$. In order to solve the operational equation η must be expressed in terms of ξ . This is done by means of the non-compatibility relation

$$(\xi\eta - \eta\xi) = \frac{i\hbar}{2\pi} [\xi, \eta].$$

The physical hypotheses arising out of the transformation may now be briefly summed up. Any normalized state vector is represented in the ξ -space by components $(\xi |)$ so that

$$\psi = \sum_{\xi} \psi(\xi') (\xi' |). \quad (38)$$

Similarly in the η -space,

$$\psi = \sum_{\eta'} \psi(\eta') (\eta' |) \quad (38')$$

so that from (36)

$$(\eta' |) = \sum (\eta' | \xi') (\xi' |), \quad (39)$$

$$\text{where } \sum (\eta' | \xi) (\xi' | \eta'') = \delta \eta' \eta''. \quad (40)$$

Similar equations follow when ξ' and η' are continuous variables. The probability that a system when referred to ξ 's is in the state ξ' is given by

$$P(\xi') = |(\xi' |)|^2. \quad (E)$$

Similarly, the probability that the system is in the same state η' is

$$P(\eta') = |(\eta' |)|^2 \quad (E)$$

when it is referred to the η 's. Moreover, on account of the reciprocal nature of the two transformation functions as given by (40) we have the theorem of reciprocity :

“The probability of the ξ 's having the values ξ' for the state for which η 's certainly have the values η' is equal to the probability of the η 's having the values η' for the state for which the ξ 's certainly have the values ξ' ”.

(Dirac : 'Q.M.' (1935) p. 66)

$$|(\xi' | \eta')|^2 = |(\eta' | \xi')|^2. \quad (F)$$

(E) leads to the definition of transition probabilities and expectation values. The vectors ψ in (38) being normalized

$$\sum_{\xi'} |(\xi' |)|^2 = 1.$$

* Here ξ' , ξ'' , η' , η'' etc. represent the eigenvalues.

This is as it should be if the probability interpretation of (E) is to remain valid. Suppose that an observable operating on $\psi(\xi)$ generates a function $\alpha \psi(\xi)$. Let

$$\alpha \psi(\xi'') = \sum \psi(\xi') (\xi' | \alpha | \xi''). \quad (41)$$

$$\text{Then} \quad \psi^\times(\xi') \alpha \psi(\xi'') = (\xi' | \alpha | \xi'') \quad (42)$$

The left-hand side is usually expressed as $(\xi' | \alpha | \xi'')$ so that

$$\alpha \psi(\xi'') = \sum \psi(\xi') (\xi' | \alpha | \xi''). \quad (43)$$

$$\text{Also,} \quad P = |(\xi' | \alpha | \xi'')|^2. \quad (44)$$

P is the probability that a system which was originally in the eigenstate $\psi(\xi'')$ goes into the state $\psi(\xi')$ as a result of the observation α being made. For a real observable

$$\psi^\times(\xi') \alpha \psi(\xi'') \text{ and } \psi^\times(\xi'') \alpha \psi(\xi')$$

are conjugate imaginaries. Hence ξ' and ξ'' may also interchange their places in (44) without changing the value of P. The expectation value for any particular state is the same as given by D.

In fine, we will consider how the expectation values⁴ change with time. Consider any operator $F(q, p)$ expressed in the q -space. For a particular state vector ψ the corresponding expectation value is

$$\bar{F} = \int \psi^\times F(q, \frac{-i\hbar}{2\pi} \partial/\partial q) \psi dq. \quad (45)$$

$$\text{Hence} \quad \frac{d\bar{F}}{dt} = \int \left[\frac{\partial \psi^\times}{\partial t} F \psi + \psi^\times F \frac{\partial \psi}{\partial t} \right] dq.$$

$$\text{Since} \quad H\psi = \frac{i\hbar}{2\pi} \frac{\partial \psi}{\partial t} \text{ and } (H\psi)^\times = -\frac{i\hbar}{2\pi} \frac{\partial \psi^\times}{\partial t},$$

$$\begin{aligned} \frac{d\bar{F}}{dt} &= \frac{2\pi i}{\hbar} \int \left[H(\psi)^\times F - \psi^\times F H \right] dq \\ &= \frac{2\pi i}{\hbar} \int \psi^\times (HF - FH) \psi dq \\ &= \int \psi^\times [F, H] \psi dq, \end{aligned} \quad (46)$$

in the notation of the Poisson bracket. Thus the mean value of the Poisson bracket $[F, H]$ equals dF/dt . This is the analogue of the classical result

$$\frac{dF}{dt} = [F, \mathcal{H}].$$

The expectation values of q 's and p 's follow the Hamiltonian equation in the limiting case when $\hbar = 0$. From (46) it is obvious that

$$\frac{d\bar{H}}{dt} = 0.$$

If we put $F = H^n$, we find also that

$$\frac{d}{dt} \bar{H}^n = 0. \quad (47)$$

As (47) is true for all values of n it gives the analogue of the principle of conservation of energy. If F is any classical integral such as that of conservation of momentum, it is true that

$$[F^n, H] = 0. \quad (48)$$

Hence
$$\frac{d}{dt} \bar{F}^n = 0 \quad (49)$$

and the quantum-mechanical analogue is presented by the constancy of the expectation value.

4. THE UNCERTAINTY PRINCIPLE

There are two aspects of the uncertainty principle; one is purely experimental while the other is theoretical. The experimental aspect is not new. It has been well-known that the most delicate operations of experimentation can disturb a system so much that an unknown measure of uncertainty vitiates the final results. The frequency theory of probability is, in fact, based upon this hypothesis of experimental uncertainty. The theoretical aspect is new, but not as new to the metaphysician as to the physicist.⁷ Consider, as an example, the hackneyed question of the position and velocity of a particle. From the experimental point of view the uncertainty principle means that both the position and the velocity cannot be accurately observed at any instant as the observation of one disturbs the other. If we take the purely theoretical standpoint the question is whether the concept of velocity at an instant has any physical meaning. Average velocity has an observational significance but 'velocity at an instant' is a logical fiction.

As the empire of known facts grows in size the physical hypotheses are contradicted and the mathematical postulates cannot be reconciled with the new logic of phenomena. Heisenberg⁵ was the first to show that the tacit assumption regarding the possibility of determining at once both the position and the velocity of a particle cannot be true in atomic physics. The corresponding modification of the theoretical structure was brought out by Born when he gave the probability interpretation of Schrödinger's wave function. Dirac's theory of incompatible observables is a full-fledged exposition of the uncertainty principle so far as it relates to quantum mechanics. The method of the second quantization for wave fields of radiation similarly establishes uncertainty relations for the measurable quantities of a radiation field.

The first intimation of the microscopic natural phenomena being governed by uncertainty relations came with the two quality relations,

$$= h\nu, \quad \hat{p} = \hbar \hat{\tau}, \quad (50)$$

which Einstein had established for light and de Broglie had assumed for matter. Consider a wave group of extent Δx associated with a particle at some instant. Let us suppose that the group is formed of wave lengths near about λ_0 . Then if n wave lengths λ_0 are contained in Δx at least $n + 1$ waves of a neighbouring wave length $\lambda_0 - \Delta\lambda$ must be contained in Δx . Hence

$$\Delta x = n\lambda_0 \geq (n + 1)(\lambda_0 - \Delta\lambda)$$

$$\text{or} \quad \frac{\Delta x \Delta \lambda}{\lambda_0^2} \geq 1.$$

$$\text{Since} \quad \tau = 1/\lambda_0 = p_0/h, \quad (1/\lambda_0^2)\Delta\lambda = \Delta p/h$$

$$\Delta x \Delta p \geq h. \quad (51)$$

We may also write this in the form

$$v\Delta t \cdot \Delta p \geq h$$

because $\Delta x = v\Delta t$, also $v\Delta p = \Delta E$ for a particle. Therefore

$$\Delta E \Delta t \geq h. \quad (52)$$

From the first relation of (50) also (52) can be deduced.

The uncertainty relation for position and momentum can also be deduced from (31). For if

$$\phi(x) = \frac{1}{(a\pi^{1/2})^{1/2}} e^{-\frac{1}{2}\left(\frac{x-x_1}{a}\right)^2 + \frac{2\pi i p_1 x}{h}} \quad (53)$$

$$\phi(p) = \frac{1}{(a\pi^{1/2})^{1/2}} e^{-\left(\frac{2\pi a}{h}\right)^2 (p-p_1)^2 - \frac{2\pi i}{h}(p-p_0)x_0} \quad (54)$$

When p has the mean value p_1 the conventional measure of uncertainty is, as given by $|\phi(p)|^2$

$$\frac{1}{\sqrt{2}} \frac{h}{2\pi a}. \quad (55)$$

Similarly when x_1 is the mean value of x the uncertainty is a . It is clear that the product of the two uncertainty measures is of the order of h . In the choice of $|\phi(x)|$ we are led to introduce a phase factor which does not alter the probability $|\phi(x)|^2$ in order that the determination of $\phi(p)$ may be simplified. This proof depends upon the form of the normalized function $\phi(x)$. But the relation (31) can also be expressed in Dirac's notation as

$$(p_0 |) = h^{-\frac{1}{2}} \int e^{-2\pi i q p_0/h} (q |) dq \quad (56)$$

If $|(q)|$ is practically negligible outside a region Δq then the variation in wave length h/p must give, by the argument which gives (51)

$$\frac{\Delta q}{h/p_0} - \frac{\Delta q}{h/(p_0 - \Delta p)} \sim 1 \quad (57)$$

or

$$\Delta q \Delta p \sim h.$$

This is brought out in Chap. IV (4) more explicitly.

A very general statement of the uncertainty relations can be deduced from the calculus of operators. The expectation value of any real observable is positive. Using this the uncertainty relations have been found out.⁶ If A and B are real observables and λ is a real number, the expectation value as shown by a bar overhead must be positive for $(A + i\lambda B)(A + i\lambda B)^*$.

$$\text{Hence} \quad \overline{(A + i\lambda B)(A + i\lambda B)^*} \geq 0. \quad (58)$$

$$\text{Therefore} \quad \overline{A^2} + \overline{B^2} \lambda^2 - i \overline{(AB - BA)} \lambda \geq 0.$$

Choosing λ so that the left-hand side becomes a minimum we have

$$\overline{A^2} - \frac{(\overline{AB - BA})^2}{\overline{B^2}} \geq 0$$

or

$$\overline{A^2} \overline{B^2} \geq \frac{1}{4} (\overline{AB - BA})^2. \quad (95)$$

Consider the observables

$$\delta A = A - \overline{A}, \quad \delta B = B - \overline{B}.$$

Then

$$\delta A \delta B = \delta B \delta A + AB - BA.$$

Hence,

$$\begin{aligned} \overline{\delta A^2 \delta B^2} &\geq \frac{1}{4} (\overline{\delta A \delta B - \delta B \delta A})^2, \\ &\geq \frac{1}{4} (\overline{AB - BA})^2. \end{aligned} \quad (60)$$

Let A and B stand for p and q then

$$\overline{AB - BA} = h/2\pi$$

and

$$(\overline{\delta p})^2 (\overline{\delta q})^2 \geq \frac{h^2}{16\pi^2}. \quad (61)$$

Modern physics presents several interesting applications of this principle. In the study of solids subjected to high pressures we are told how a limitation on the closeness of packing arises and why solid matter would assume an amorphous form. In the theory of spectra the principle explains the natural line breadth. The phenomenon of the fluctuation of the electric force vector, when there are no photons in the field, is also understood in the light of the uncertainty relations. In quantum statistics, for the same reason, the area of the smallest cell in the p - q space is h . We will close this account with a discussion of the angular momentum vector, (H_x, H_y, H_z) .

We have for a particle in a central field of force

$$\begin{aligned} yP_z - zP_y &= H_x, \\ zP_x - xP_z &= H_y, \\ xP_y - yP_x &= H_z. \end{aligned}$$

If $\frac{2\pi}{h} (H_x, H_y, H_z) = (M_x, M_y, M_z)$

the operational equations become

$$\begin{aligned} y \frac{\partial \psi}{\partial z} - z \frac{\partial \psi}{\partial y} &= i M_x \psi, \\ z \frac{\partial \psi}{\partial x} - x \frac{\partial \psi}{\partial z} &= i M_y \psi, \\ x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} &= i M_z \psi. \end{aligned} \quad (62)$$

These three equations are not compatible with each other unless $M_x = M_y = M_z = 0$. Hence it is not possible to know simultaneously all the three components of the angular momentum vector. They are not a set of compatible observables. It can be shown by direct computation that

$$\begin{aligned} H_x H_y - H_y H_x &= (y P_z - z P_y) (z P_x - x P_z) - (z P_x - x P_z) (y P_z - z P_y) \\ &= y P_z z P_x - z^2 P_y P_x - yx P_z^2 + zx P_y P_z \\ &\quad - zy P_x P_z + xy P_z^2 + z^2 P_x P_y - x P_z z P_y \\ &= y P_x (P_z z - z P_z) - x P_y (P_z z - z P_z) \\ &= (y P_x - x P_y) (P_z z - z P_z) \\ &= \frac{h}{2\pi i} P_z. \end{aligned} \quad (63)$$

$$\text{But } H^2 H_x = H_x H^2, H^2 H_x = H_x H^2, H^2 H_y = H_y H^2. \quad (64)$$

Hence H^2 and H_x can be regarded as compatible. So it is possible to know the magnitude of the resultant and also of one of the components simultaneously. The uncertainty principle tells us why the maximum possible value of a component does not agree with that of H . The reader is reminded of the fact that if the maximum value of H_x is $Jh/2\pi$, the value of H^2 is $J(J+1)h^2/4\pi^2$. If the maximum were the same as the value of H the knowledge of one component would be enough to know the vector, the other components being zero. Thus all the three components H_x, H_y, H_z being known, by Heisenberg's uncertainty relations, the angles about the x, y, z axes would be completely undetermined. This is only possible if the probability distribution is spherically symmetrical, that is, if the angular momentum is zero.

(G. Herzberg, "Atomic Spectra and Structure" (1937) p. 102).

5. THE ELECTRON SPIN

The theory of generalized operators outlined above has been successfully applied to explain atomic and molecular spectra and to various problems of physics and chemistry. The methods have also attained considerable success

in nuclear physics. The statistics of charged particles, the arrangement of atoms in crystals, the distributions of the electronic shells about the nucleus and spins of the valence electrons: these are questions on which quantum mechanics has given accurate answers as verified in the applications to ferromagnetism, conductivity, valence, etc. It is the electron spin that happens to play the most conspicuous part in all these phenomena. The spin hypothesis was first suggested by Compton in 1921 and later by Goudsmit and Uhlenbeck in 1925. It led immediately to an explanation of the multiplets, the anomalous Zeeman effect and the Paschen-Back effect as well. The Pauli principle was formulated and it cleared up at once most of the long-standing difficulties of the periodic table. The mystery of the Landé g -formula was partially solved but the factor 2 which arises in the relationship between the electronic magnetic moment and the angular momentum continued to be baffling in spite of the explanation given by L. H. Thomas. Dirac's relativistic theory of the electron gave a logical explanation of the spin and resolved the difficulty of the factor 2.

From the relativistic form³ of the classical Hamiltonian we get for the wave equation of a particle,

$$\left\{ W/c - (m^2 c^2 + p_x^2 + p_y^2 + p_z^2)^{\frac{1}{2}} \right\} \psi = 0, \quad (65)$$

in the usual notation. This equation merely takes account of the relativistic variation of mass but has not the proper symmetry between the energy and momentum components to give the necessary invariance. By multiplying the operator on ψ in (65) by another, viz.,

$$\frac{W}{c} + (m^2 c^2 + p_x^2 + p_y^2 + p_z^2)^{\frac{1}{2}}$$

it is possible to get a symmetric wave equation,

$$\left\{ W^2/c^2 - (m^2 c^2 + p_x^2 + p_y^2 + p_z^2) \right\} \psi = 0. \quad (66)$$

The last equation is not permissible on the postulates of quantum mechanics as the energy operator is not linear. Hence Dirac undertook to investigate whether a linear equation of the form

$$\left\{ W/c + \alpha_x p_x + \alpha_y p_y + \alpha_z p_z + \beta \right\} \psi = 0, \quad (67)$$

invariant under the Lorentz transformation, cannot be true for an electron in free space. The operators α 's and β must be independent of the energy and momentum and also of the co-ordinates x, y, z, t . Hence they must bring out some innate property of the electron such as the spin, if it exists. Multiplying (67) by $\left\{ W/c - \alpha_x p_x - \alpha_y p_y - \alpha_z p_z + \beta \right\}$ we get

$$\left\{ \frac{W^2}{c^2} - \sum_{xyz} [\alpha_x^2 p_x^2 + (\alpha_x \alpha_y + \alpha_y \alpha_x) p_x p_y + (\alpha_x \beta + \beta \alpha_x) p_x] \right\} - \beta^2 \psi = 0. \quad (68)$$

Comparing (68) with (66) we find

$$\alpha_x^2 = 1, \alpha_x \alpha_y + \alpha_y \alpha_x = 0, \alpha_x \beta + \beta \alpha_x = 0, \quad (69)$$

and similar relations for the other components of α . One effect of the comparison with (66), is that the new equation has to admit negative values of the energy. We can obtain the following representation of α :

$$\alpha_x = \rho_1 \sigma_x, \alpha_y = \rho_1 \sigma_y, \alpha_z = \rho_1 \sigma_z, \alpha_m = \rho_3 \quad (70)$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$\rho_1 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad \rho_2 = \begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix}, \quad \rho_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The wave equation (67) may now be written in the form

$$\left\{ W/c + \rho_1 (\hat{\sigma} \cdot \hat{p}) + \rho_3 mc \right\} \psi = 0. \quad (71)$$

Hence following the classical procedure of replacing W and \hat{p} by $W + e \phi$ and $\hat{p} + e \hat{A}/c$ for the motion in an electromagnetic field, we write

$$\left\{ W/c + \frac{e\phi}{c} + \rho_1 (\hat{\sigma} \cdot \hat{p} + \frac{e}{c} \hat{A}) + \rho_3 mc \right\} \psi = 0. \quad (72)$$

If the four-vector \hat{A}_μ stands for ϕ, A_x, A_y, A_z and the four-vector \hat{p}_μ for $W/c, p_x, p_y, p_z$ we can write (72) in the symmetrical form*

$$\left\{ \alpha_\mu (\hat{p}_\mu + e/c A_\mu) + \alpha_m mc \right\} \psi = 0. \quad (73)$$

The Lorentz invariance of (73) can be proved by showing that if

$$\hat{p}_\mu = a_{\mu\nu} \hat{p}'_\nu, \quad A_\mu = a_{\mu\nu} A'_\nu, \quad \psi' = \gamma \psi \quad (74)$$

$$\text{where} \quad \gamma^\times \alpha_\nu \gamma = \alpha_\mu a_{\mu\nu}, \quad \gamma^\times \alpha_m \gamma = \alpha_m \quad (75)$$

the equation (73) transforms into

$$\gamma^\times \left\{ \alpha_\nu (\hat{p}'_\nu + e/c A'_\nu) + \alpha_m mc \right\} \psi' = 0. \quad (76)$$

When the equation (72) is written in the quadratic form, on multiplying it by the factor

$$\frac{W}{c} + \frac{e\phi}{c} - \rho_1 (\hat{\sigma} \cdot \hat{p} + \frac{e}{c} \hat{A}) - \rho_3 mc, \quad (77)$$

* The dummy-suffix convention has been used in (73)-(76).

we compare it with the classical quadratic equation obtained from (66) after making the above mentioned alterations for an electromagnetic field ; we find that the new equation is equivalent to

$$\left\{ \left(\frac{W}{c} + \frac{e}{c} \phi \right)^2 - \left(\hat{p} + \frac{e}{c} \hat{A} \right)^2 - m^2 c^2 - \frac{he}{2\pi c} (\hat{g}, \hat{H}') - i \rho_1 \frac{he}{2\pi c} (\hat{g}, \hat{E}') \right\} \psi = 0 \quad (78)$$

if the form of the classical Hamiltonian is resorted to. The new terms are the last two. In them \hat{H}' , \hat{E}' , are the magnetic and electric intensities. Because calculation shows

$$\hat{H}' = \text{curl } \hat{A}, \hat{E}' = - \frac{1}{c} \left(\frac{\partial \hat{A}}{\partial t} + \text{grad } \phi \right). \quad (79)$$

Hence the electron possesses the additional potential energy,

$$\frac{he}{4\pi mc} (\hat{g}, \hat{H}') + i \rho_1 \frac{he}{4\pi mc} (\hat{g}, \hat{E}')$$

unknown in the classical treatment. The electron has a magnetic moment $-he/4\pi mc \hat{g}$ and an electric moment $-i \rho_1 he/4\pi mc \hat{g}$ which is imaginary. When we calculate the angular momentum by the formulae

$$M_x = y p_z - z p_y, M_y = z p_x - x p_z, M_z = x p_y - y p_x$$

the electron is found to have a spin angular momentum $\hbar/4\pi$. "The same vector \hat{g} fixes the directions of both the spin magnetic moment and the spin angular momentum. If an electron in a certain state of spin has a spin angular momentum $\frac{1}{2} \hbar$ in a particular direction, it will have a magnetic moment $-eh/2mc$ in the same direction."

(Dirac "Q.M." (1935) p. 265)

Dirac's \hbar is equal to $h/2\pi$ in our notation.

REFERENCES

- (1) CONDON AND MORSE, "Quantum Mechanics" pp. 187-215, (1930).
- (2) FRENKEL, J., "Wave Mechanics," part 2, pp. 162-76 (1934).
- (3) DIRAC, P. A. M., "Quantum Mechanics," pp. 250-65 (1935).
- (4) TOLMAN, R. C. "The Principles of Statistical Mechanics" pp. 237-43 (1938).
- (5) HEISENBERG, W., "The Physical Principles of the Quantum Theory", pp. 13-54 (1930).
- (6) BORN, M., "Atomic Physics" p. 312 (1935).
- (7) FLINT, H. T., "Reports on Progress in Physics," 4, pp. 163-72 (1937).

VII THE ATOM

1. The Problem of n Particles. 2. The Energy Levels in the Atom. 3. The Number of Elements.

1. THE PROBLEM OF n PARTICLES

Consider a system of n identical particles ; they may be electrons, neutrons or atomic nuclei. It is not possible to distinguish one such particle from another of the same group. This fact is often referred to as the identity principle. Hence, if two such particles interchange their places the physical system does not alter, which has to be taken account of, in a study of the physical states of two or more identical particles. In a theoretical presentation one associates suffixes and coordinates with particles for the sake of convenience and clarity. But they prove to be distinguishing marks and therefore the description of a physical state has to be so given that the different suffixes and coordinates describe only one situation even if the electrons represented by them interchange places. Suppose¹ that f is an observable of the system. The eigenvalues of f must be the same as those of Pf , if P is any permutation of electrons. The restriction on the possible observable f of the system is sufficient to obtain all information about it. Here the equation of expectation values in a state ψ is, in the usual notation,

$$\int \psi^* f \psi d\tau = \int \psi^* (Pf) \psi d\tau. \quad (1)$$

If the operation P^{-1} is performed on each of the functions in the second integral the latter must still be unaltered because of the identity principle. Therefore,

$$\int \psi^* f \psi d\tau = \int (P^{-1}\psi)^* f (P^{-1}\psi) d\tau, \quad (2)$$

$$\text{or} \quad \int \psi^* f \psi d\tau = \int \psi_0^* f \psi_0 d\tau, \quad (3)$$

$$\text{where} \quad P^{-1}\psi = \psi_0. \quad (4)$$

If (3) is to hold good for all observables it is just sufficient to take

$$f = \delta(x - a_1) \delta(x_2 - a_2) \dots \dots \dots$$

$$\text{to show that} \quad \psi^* \psi = \psi_0^* \psi_0 \quad (5)$$

$$\text{at all points and} \quad f = \delta(x_1 - a_1) \delta(x_2 - a_2) \dots \dots \dots \delta(x_i - a_i)$$

to get the other relation,

$$\psi^* \frac{\partial \psi}{\partial x_i} = \psi_0^* \frac{\partial \psi_0}{\partial x_i}, \quad (6)$$

at all points. From (5) and (6) it follows that

$$\psi = \psi_0 e^{i\alpha} = P^{-1} \psi e^{i\alpha},$$

$$\text{or} \quad P\psi = \psi e^{i\alpha} \quad (7)$$

α being a constant. Consider now any ψ function described as $\psi_{ijk} \dots$ where each suffix stands for an electron. Let an interchange of electrons i, j take place so that

$$\psi_{jik} \dots = \psi_{ijk} \dots e^{i\alpha}.$$

If again an interchange of the same two electrons take place we have

$$\psi_{ijk} \dots = \psi_{jik} \dots e^{i\alpha} = \psi_{ijk} \dots e^{2i\alpha}.$$

$$\text{Hence} \quad e^{2i\alpha} = 1, \quad e^{i\alpha} = \pm 1. \quad (8)$$

From the probability formula, $\psi^* \psi$, it follows that the effect of a permutation may be of the perfectly general form (7) as the phase does not affect the probability density. However the only functions which are of physical interest are given by (8). If

$$\psi_{ijk} \dots = -\psi_{jik} \dots \quad (9)$$

the system is said to be antisymmetrical while the wave function which obeys the symmetry condition,

$$\psi_{ijk} \dots = \psi_{jik} \dots \quad (10)$$

gives the symmetrical system. The symmetry or antisymmetry of a system is a physical quality and it is independent of the coordinate system used. Dirac⁶ has shown from the laws of quantum mechanics that the symmetric or antisymmetric nature of a system remains invariant for canonical transformations. From the laws of motion he has also deduced that a system which is at one time symmetrical (or antisymmetrical) remains always symmetrical (or antisymmetrical). This particular property may be verified from the familiar Schrödinger equation for electrons expressed in Cartesian coordinates. The equation runs as follows :

$$\frac{h}{2\pi i} \frac{\partial \psi}{\partial t} = \left[\frac{h^2}{8\pi^2 m} (\nabla_1^2 + \nabla_2^2 + \dots + \nabla_n^2) - V \right]. \quad (11)$$

The potential energy is a symmetric function of the coordinates. Hence if ψ is symmetrical or antisymmetrical the same property is shared by, $\partial \psi / \partial t$. This shows the invariance of the particular characteristic enjoyed by ψ .

Let us discuss at some length the field inside a many-electron atom. ~~The mathematical discussion for any assembly of n electrons can be carried~~

out on precisely the same lines. If the nucleus has a charge ze and the number of electrons is n the Hamiltonian is

$$H = \sum_{i=1}^n \left(\frac{1}{2m} \hat{p}_i^2 - \frac{ze^2}{r_i} + \xi(r_i) \hat{L}_i \cdot \hat{S}_i \right) + \sum_{i>j=1}^n \frac{e^2}{r_{ij}}. \quad (12)$$

The motion of the nucleus is assumed to play a negligible part. The distances of the electrons from the nucleus are given by r_i and from each other by r_{ij} . The spin-orbit interaction energy is assumed to be $\xi(r_i) \hat{L}_i \cdot \hat{S}_i$. Thomas and Frenkel have shown that the spin-orbit interaction energy for an electron in a central field with potential $V(r)$ is

$$\frac{1}{2m^2 c^2} \left(\frac{1}{r} \frac{dV(r)}{dr} \right) \hat{L} \cdot \hat{S} \equiv \xi(r) \hat{L} \cdot \hat{S} \quad (13)$$

where \hat{L} is the orbital and \hat{S} the spin angular momentum. In writing (12) we assume that this form of coupling gives a very good approximation to the real interaction energy. When the i th electron is at a great distance from the nucleus as compared to the other electrons, r_{ij} is approximately the same as r_i for the $n-1$ distances. Hence the effect of the mutual repulsion is to produce a screening so that the nuclear potential, as toned down by the last term of (12), becomes $-(z-n+1)e^2/r_i$ for the i th electron. It is now possible to split up (12) into two parts, $H = H_0 + H'$:

$$H_0 = \sum_i \left[\frac{1}{2m} \hat{p}_i^2 + V(r_i) \right], \quad (14)$$

$$H' = \sum_i \left[\xi(r_i) \hat{L}_i \cdot \hat{S}_i - \frac{ze^2}{r_i} - V(r_i) \right] + \sum_{i>j=1}^n \frac{e^2}{r_{ij}}. \quad (15)$$

The Schrödinger equation for (14), viz.,

$$H_0 \psi = E_0 \psi \quad (16)$$

possesses solutions of the form

$$\psi = \psi_1(\theta_1) \psi_2(\theta_2) \dots \psi_n(\theta_n) \quad (17)$$

where θ_r stands for the four quantum numbers, n, l, m, s of the r th electron. Since ψ_1, ψ_2 etc. all satisfy the same differential equation, except for the eigenvalue constant which is different for different functions, $\psi_1 = \psi_2$ when $\theta_1 = \theta_2$. Each electron moves in a central field as if the other electrons are absent:

$$\left[\frac{-\hbar^2}{2m} \nabla^2 + V(r) \right] \psi(\theta_i) = E_0(\theta_i) \psi(\theta_i); \quad (18)$$

$$E_0 = \sum_{i=1}^n E_i(\theta_i). \quad (19)$$

There is a two-fold degeneracy of ψ . Firstly $\psi_r(\theta_r)$ is independent of m and s . Secondly, the total energy E_0 is the same for the $n!$ different permutations of the electrons carried out through the suffixes. The most general solution of (16) is therefore of the form

$$\psi = \sum_{i=1}^{n!} C_i P_i \left[\psi_1(\theta_1) \psi_2(\theta_2) \cdots \psi_n(\theta_n) \right] \quad (20)$$

where each P_i is a different substitution, being one of the $n!$ substitutions on n letters. C 's are all consonants. If the electrons constitute a symmetrical system,

$$\text{all } C\text{'s are equal.} \quad (21)$$

If they constitute an unsymmetrical system,

$$C_i = (-1)^i A, \quad (22)$$

where A is a constant and i is the number of transpositions into which P_i can be resolved by any method. Let us take $\psi_1, \psi_2 \dots$ etc as normalised wave functions. For a system of n particles we can have two possible ψ functions.

$$\psi = \frac{1}{(n!)^{1/2}} \begin{vmatrix} \psi_1(\theta_1) & \psi_1(\theta_2) & \cdots & \psi_1(\theta_n) \\ \psi_2(\theta_1) & \psi_2(\theta_2) & \cdots & \psi_2(\theta_n) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_n(\theta_1) & \psi_n(\theta_2) & \cdots & \psi_n(\theta_n) \end{vmatrix} \quad (23)$$

The square array is to be evaluated as a determinant for an antisymmetrical system and as a permanent (which has all the terms of the determinant written with the same sign) for a symmetrical one. The electrons in an atom constitute an anti-symmetrical system. For according to the Pauli exclusion principle no two electrons in the atom are found to be with the same set of quantum numbers $\theta_i = \theta_j$; and when we put any two θ 's as equal in the determinant (23) it vanishes as two columns become identical. Hence the non-existence of electrons with the same quantum numbers is explained on the hypothesis that the electrons in an atom are represented by an antisymmetrical ψ . Quantum mechanics has nothing to say as to why the Pauli principle holds good in the atom. We merely draw the conclusion from this observed fact that electrons satisfy the Fermi-Dirac statistics or that their statistical states are governed by an anti-symmetrical wave function. Similarly there is no quantum-mechanical explanation why the protons obey the Einstein-Bose statistics, that is, why the Schrödinger function of the orchestra of harmonic oscillators for radiation is symmetrical. The law of black-body radiation merely tells us that Nature has favoured this statistics for radiation.

An examination of the rotational bands of diatomic molecules reveals what statistics the nuclei follow. The ratio of the intensities of alternate

bands gives the nuclear spin ; what lines are weak and which are strong is decided by the statistics that the nuclei obey. There is ample evidence that the nitrogen nucleus obeys the Einstein-Bose statistics. This would not be reconciled with the earlier view of the constitution of that nucleus. Because electrons and protons both follow the other statistics ; and if there were 14 protons and seven electrons in the nitrogen nucleus it would claim an antisymmetrical ψ . For example consider two atomic nuclei each of which consists of m electrons, p protons and n neutrons³. Neutrons also satisfy the Fermi-Dirac statistics. Hence one interchange of two electrons, of two protons or of two neutrons changes the sign of ψ once. If the two nuclei are to exchange places it means $m + n + p$ exchanges for the constituent particles. The sign of ψ would be altered into

$$(-1)^{m+n+p}$$

in this exchange of two nuclei. When $m + n + p$ is even the Einstein-Bose statistics is obeyed ; when $m + n + p$ is odd the other statistics is valid. If the old view of the constitution of the nitrogen nucleus were correct, $m = 7$, $n = 0$, $p = 4$ and the answer would be wrong. But according to the new view which has many other arguments in support, $m = 0$, $n = 7$, $p = 7$ and the right answer is obtained.

2. THE ENERGY LEVELS IN THE ATOM

To a certain degree of approximation the electrons in an atom may be assumed to be characterized by four quantum numbers in the manner explained previously². Each set of spatial quantum numbers gives a level which by the Pauli principle, is occupied by two electrons at the most. When there are two they are paired, that is, their spins are $\pm\frac{1}{2}$. In a heavy atom which is placed near uranium in the periodic table, the various levels of energy in the ascending order are :

$$1s; 2s, 2p; 3s, 3p, 3d, 4s, 4p, 4d, 4f; 5s, 5p, 5d; 6s, \dots \quad (A)$$

On the other hand, the levels in a lighter atom with atomic number less than 20, are arranged in the order :

$$1s; 2s, 2p; 3s, 3p; 4s; 3d; 4p; 5s; 4d; 5p; 6s; 5d; 4f; \dots \quad (B)$$

There is a gradual transition from the latter order to the first as we pass on from calcium to the higher atoms. The atom should be looked upon as a hollow in the usual energy representation, the various discrete levels of negative energy ranging from the lowest 1s level to the top beyond which there is a continuous range of positive energy levels due to hyperbolic orbits. The break in the order as in (B) is due to the interaction energy being too large. The various elements are grouped in the periodic table according to the character of the electrons in the outer shell. The inert gas column is characterised by closed shells and $s^2 p^6$ subshells. The hydrogen column contains

elements with only an s electron in the outershell. It is interesting to see how potassium is promoted to this column through the sudden break in order at $3p$ as shown in (B). The alkaline earths are characterized by two s electrons in the outermost shell. Since the periodic table is mainly concerned with the grouping of elements according to their chemical properties it is but to be expected that the outermost or valence electrons should be the chief index for position in the table. It must be emphasized however that the physical properties such as conductivity and refractive index vary from element to element in the same column as they are not dependent exclusively on the outer electrons. The central principle on which the different levels of the normal atom (in an unexcited or ground state) are occupied in what is called the building-up principle. According to it the electrons occupy the lowest possible levels without violating the exclusion principle.

It is by means of spectra that the theory of the order of levels is verified. With each atom is associated a term type of the ground state as also of the other states. From the spectral character these types can be inferred. On calculating the term types theoretically it is ascertained that the grouping of electrons in the shells as given is correct. How the inner $3d$ shell is gradually filled up in the case of the eight elements Sc to Ni, the inner $4d$ shell for the group of eight, Y to Pd, the inner $4f$ shell for the rare earth and the $3d$ shell for the atomic numbers (71-78)—all these facts are suggested by the term types and the variation of the ionization potential from element to element. Some preliminary explanations will have to be given before the calculation of the term type is made clear.

The angular momentum for the spin of the electron is always $h/4\pi$, the magnitude of the spin itself being $\frac{1}{2}$. We will take $h/2\pi$ as the unit of angular momentum for the sake of convenience. When there are n electrons in the atom their spin vectors \hat{S}_i combine to give the resultant spin vector whose allowed values are

$$\frac{n}{2}, \frac{n}{2} - 1, \dots, \frac{1}{2} \text{ or } 0.$$

The series terminates with 0 if n is even and with $\frac{1}{2}$ if n is odd. When all the spin vectors are parallel we get $n/2$. When the allowed value is S , the exact amount of the angular momentum (scalar) is $\sqrt{S(S+1)}$ according to the rigorous treatment of quantum mechanics. Obviously S is integral (or half-integral) if the number of electrons is even (or odd). The multiplicity of an atomic term is defined as $2S + 1$. The term is even (or odd) as the number of electrons is odd (or even). We now understand why the atoms arranged in the order of their atomic number have alternately even and odd term types. The reason why $2S + 1$ is called the multiplicity may also be made clear now.

Let the resultant orbital angular momentum vector for the electrons be \hat{L} with the maximum value given by the integer L . Then the law of interaction

between \hat{L} and \hat{S} permits for $L > S$, the following $2S+1$ values for the resultant :

$$L - S, L - S + 1, \dots L + S.$$

But if $S > L$, the following $2L+1$ values are permitted :

$$S - L, S - L + 1, \dots S + L.$$

There are therefore $(2S+1)$ different values of energy in the more frequent case $L > S$. Hence $(2S+1)$ is defined as the multiplicity of the term. When we speak of the 3S term for which $L=0$ and $S=1$, there being two electrons, we describe the multiplicity as 3 just to stick to this convention although the number of distinct energy levels is $2L+1=1$ in this case. The resultant of the \hat{L} , \hat{S} vectors is denoted by \hat{J} .

To determine the vector \hat{L} it is not necessary to take account of all the electrons in the atom. Those in closed shells and subshells may be dropped. For $\sum m = 0$ for all the levels of a subshell, $hm/2\pi$ being the component of the orbital angular momentum in an arbitrary direction. For the calculation of \hat{L} and obviously for that of \hat{S} as well it is sufficient to consider the electrons of the outermost open subshell or subshells. In the case of two electrons with orbital angular momenta \hat{l}_1 and \hat{l}_2 the possible values of the resultant are

$$l_1 + l_2, l_1 + l_2 - 1, \dots |l_1 - l_2|.$$

Successive additions of vectors are to be carried out by this rule.

In the case of a small interaction between the electrons it is represented by a small velocity of precession of the vectors $\hat{l}_1, \hat{l}_2, \dots$ about the resultant \hat{L} and the orbital velocities have a physical meaning. But when the velocity of precession is large the interaction is strong and the orbital velocities cease to have any meaning. The mathematical procedure, however, allows the use of the individual quantum numbers even in the case of strong interaction. This is because the content of Ehrenfest's adiabatic principle is implicit in the assumptions about ϕ . When there is a strong coupling between \hat{L} and \hat{S} the vectors appear to possess a large velocity of precession about the resultant vector \hat{J} . When $L=0$ the term is S ; for $L=1$ it is a P term; $L=2$ gives a D term and so on. Thus when we say that the ground state of oxygen is 3P_2 we mean that $L=1, S=1$. That accounts for the symbol P and the prefix 3. The total angular momentum J is given by the suffix 2. There are 3 different J values of which the one that corresponds to the lowest energy must be specified and that is done in this case by 2. A term is said to be normal or regular when the energy values of the term increase with J . But if they decrease the term is called inverted. When the number of outer electrons is more than half of what the subshell can contain as in the case of chlorine, the inversion takes place. The result is dependent upon detailed calculations of the spin-orbit interaction matrix. The actual derivation of

the term symbol which depends upon the value of L is not a simple matter. Hund's rule is found to be successful here. It says, "Of the terms given by equivalent electrons, those with greatest multiplicity lie deepest, and of those the lowest is that with the greatest L " (Herzberg, "Atomic spectra and structure" (1927) p. 135).

The anomalous Zeeman effect can be calculated quite accurately without going into the intricacies of the matrix mechanics. The resultant magnetic moment is to be calculated for the atom on the assumption (proved by Dirac's relativistic treatment of the electron) that the ratio of the magnetic moment to the angular momentum is twice (for the spin) of that for the orbital motion. The Paschen-Back effect in strong magnetic fields is explained by the fact that the $\hat{L}\cdot\hat{S}$ (Russell Saunders) coupling which we have tacitly assumed above ceases to give a good approximation to the interaction energy and that the $\hat{J}\cdot\hat{J}$ coupling scheme has to be assumed, \hat{J} being the total angular momentum for an electron.

The existence of two kinds of helium is now very easily understood. Parhelium exhibits transitions between singlet states $S = 0$ and orthohelium the triplet states $S = 1$. When the electron spins in helium are paired we get $S = 0$ and the spatial quantum numbers satisfy the condition of symmetry. But if the two electrons are parallel or 'bachelors' $S = 1$ and there is an antisymmetry of the ψ -function for the spatial co-ordinates. To the first degree of approximation the selection rule, $\Delta S = 0$ holds good for the emission lines. But it breaks down for the higher atoms. The appearance of the so called forbidden lines does not imply that the theory is wrong; on the contrary it supports the theory inasmuch as weak lines which are not permitted by the selection rules, derived from first order considerations, are given by the higher approximations. The concept of a metastable state is dependent upon this notion of forbidden lines. In the case of helium 2^3S is a metastable state and it is above the 1^1S level, because the transition from the former to the latter is impossible as first order considerations show.

We have already referred to the possibility of forbidden lines⁴ arising from the quadrupole moment. The integrals giving the transition probabilities do not identically vanish for the quadrupole moment if they do for the dipole moment. Secondly, in a stationary state when \hat{L} and \hat{S} differ from zero the total magnetic moment vector does not lie in the same direction as the angular momentum vector on account of the notorious factor 2 in connection with the spin. The magnetic moment therefore precesses about \hat{J} being ultimately responsible for forbidden lines given by the magnetic dipole radiation.

3. THE NUMBER OF ELEMENTS

The ancient Hindus and the Greeks believed that the material universe is built up of five elements. They were not far from the modern view if

they meant by the elements the three states of matter, solid, liquid and gaseous and ether and radiation. It was too broad a hypothesis to give birth to a new science. What is meant by elements in physics or chemistry to-day is quite different from what the ancients meant by the five elements.

All substances are complex or simple. A complex substance can be shown to be made up of other substances with different chemical properties; a simple substance cannot be further analysed into chemically different substances and is called an element. Table salt, water and sugar are illustrations of complex substances while carbon, copper and sulphur are simple substances of elements. From prehistoric times are known the elements, carbon, sulphur, iron, copper, silver, tin, gold and lead. Until recently only ninety different elements were known. They have now discovered also ekaiodine and ekacæsium which are very similar to iodine and cæsium in their chemical properties. Their respective numbers in the periodic table are 85 and 87 respectively.

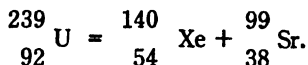
What is the maximum possible number of chemical elements? The adjective 'chemical' is used because it is the chemical properties that distinguish one element from another. There may be any number of distinct isotopes of an element, all of them being chemically similar but varying in their physical properties. Thus there are three isotopes of hydrogen already known; their weights are approximately in the proportion 1 : 2 : 3. Jeans, in his theory of the internal constitution of stars, had come to the conclusion that stellar stability depends upon the existence of elements of atomic number nearly 95. It was advanced that the elements from hydrogen to uranium, the first ninety-two, abound in the surface while deep down in the interior of a star only heavier elements exist. Although no element heavier than uranium had been discovered the fact that the energy expression of the relativistic Bohr model becomes infinite for the atomic number Z , nearly equal to 137, was interpreted as meaning that the maximum number of chemical elements is 137. The question apart of the possible existence of transuranic elements Jeans's theory of the stability of stars is no longer valid.

Discussing the question in 1928 Flint and Richardson obtained a lower limit on the highest atomic number. Their limit was 98. Even in this case the limit was found as a restriction put by relativity on the existence of a minimum radius for a circular orbit in the Bohr atom.

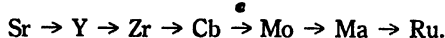
The author took up the question of the highest atomic number in 1932 while working under Eddington on the latter's relativistic theory of the interaction of two electrons. It was discovered that out of possible 136 rotations only 91 left the interaction term unchanged. This was supposed to mean that the highest atomic number is 92 and that transuranic elements do not exist.

In 1934 however came the discovery in Fermi's laboratory of what looked like transuranic elements, numbers 93 and 94. Uranium was bombarded with neutrons. The nuclei became unstable with the capture of neutrons. But

they appeared to become stable with the emission of electrons. This process suggested the existence of transuranic elements. Unstable uranium gave rise to four different types of emanations with half-periods 10 sec., 40 sec., 13 min. and 90 min. As there are only three isotopes of uranium one of these periods was supposed to be due to the formation of element 93. If it existed it was a homologue of Mn and therefore if any manganese salt was added to the uranium mixture the element 93 was expected to go down with the precipitate MnO_2 . Such a precipitate was actually discovered, diminishing by about 15 per cent. the 13-minute and 90-minute activities. The precipitate showed no trace of U or UX_1 or UX_2 . These facts were interpreted as evidence of the existence of element 93. It was argued by Noddack that the chemical evidence did not necessarily suggest Fermi's conclusion but the alternative hypothesis of uranium breaking up into radioactive isotopes of smaller atoms was not generally accepted. Series of chemical tests were devised in later years by Hahn and Meitner, by Curie and Savitch and by Hahn and Stressmann for the examination of active products of bombarded uranium. The experiments were responsible for the sensational conclusion that Noddack was right and that what appeared like transuranic elements are actually radio-active isotopes of lighter atoms obtained by fission. The fission⁵ of a uranium nucleus may be described as follows :



The $\begin{array}{c} 140 \\ 54 \end{array} \text{Xe}$ nucleus undergoes four successive beta-disintegrations producing in turn $\text{Cs} \rightarrow \text{Ba} \rightarrow \text{La} \rightarrow \text{Ce}$. Similarly, the other nucleus undergoes six beta-disintegrations according to the chain



The energy that evolves in the course of the fission and disintegrations is obtained by Einstein's relation,

$$\delta E = c^2 (m_0 - m_1 - m_2)$$

where m_0 is the mass of the original U atom and m_1 and m_2 are the masses of the final products Ce and Ru, c being the velocity of light. Similar fission has been observed for Th and Pa.

The phenomenon of nuclear fission has given rise to various questions. How is it that a U-nucleus which is stable gets broken up so easily by a neutron? Why does the nucleus break unsymmetrically? Why is it that no nucleus of mass greater than 238 is found in nature? Moreover, as it is found theoretically that energy can be gained by splitting up nuclei of weight greater than 110 it is a mystery how heavier nuclei exist in a very stable form.

The theoretical questions have been discussed by Bohr, Wheeler and others and Bohr's view is becoming popular that the nucleus is something

like a spherical drop of fluid, its energy being partly in the surface and partly in the volume. Two kinds of forces are supposed to be at work : the forces of attraction which are of the van der Waals type and are rooted in the surface ; the forces of repulsion whose seat of action is spread over the volume. The nucleus is supposed to be stable so long as the spherical shape is one of minimum total energy. As the sphere increases in size there comes a stage when the sphere ceases to be the shape for minimum energy. It is believed that this stage is reached in U. A more quantitative theory will not only answer the question mooted above but the problem of the highest atomic number will definitely be solved thereby. All recent work shows that transuranic elements most probably do not exist.

But we can no longer say that the ninety-two elements, chemically distinct as they are, are physically independent. Transmutation of elements was begun by Rutherford with α particles and he got oxygen out of nitrogen in his first successful attempt. The discovery of neutrons by Chadwick and the construction of the cyclotron by Lawrence have considerably accelerated the progress in this direction. The latest is the fission of the very heavy nuclei described above.

In the beginning of the section substances were described as simple or complex. The nuclear particles such as electrons, neutrons and the rest are neither simple nor complex substances. They possess rest masses and must therefore be distinguished from electromagnetic radiations. It is safe to look upon these particles as substances of nuclear chemistry.

Until very recently it was generally believed that the transmutation of hydrogen into heavier elements takes place in the stars. Latest studies of the sources of stellar energy show that hydrogen is being steadily transformed into helium mainly through the agency of autocatalization in most of the stars and the building up of heavy elements out of hydrogen is no longer supposed to be possible at the temperature-pressure conditions obtaining in the interior of stars.

REFERENCES

1. CONDON AND SHORTLEY, "*The Theory of the Atomic Spectra*," pp. 158-69 (1935).
2. HERZBERG, G., "*The Atomic Spectra and the Atomic Structure*" pp. 82-111. 1930-34 (1937).
3. FRENKEL, J., "*Wave Mechanics*" Part I, pp. 172-87 (1936).
4. BOWEN, I. S., "*Rev. Mod. Phys.*" 8, 2, pp. 55-81 (1936).
5. TURNER L. A., "*Rev. Mod. Phys.*" 12, 1, pp. 1-29 (1940).
6. DIRAC, P. A. M., "*The Principles of quantum Mechanics*" pp. 209-28 (1935).
7. NARLIKAR, V. V. "*The Highest Atomic Number*" from "*Nature*" March 12, (1939).
8. BETHE, H. A. "*Phys. Rev.*," 55, pp. 434-56 (1939).

VIII RADIATION

1. The Classical Theory of Radiation. 2. The Quantum Theory of Radiation.
3. Phenomenological Successes. 4. Peculiar Features and Difficulties. 5. The
Positive Electron.

1. THE CLASSICAL THEORY OF RADIATION

The classical theory¹ of radiation is based on the following field equations :

$$\text{curl } \hat{E} + \frac{1}{c} \frac{\partial \hat{H}}{\partial t} = 0, \quad (1a)$$

$$\text{div } \hat{H} = 0, \quad (1b)$$

$$\text{curl } \hat{H} - \frac{1}{c} \frac{\partial \hat{E}}{\partial t} = \frac{4\pi}{c} \rho \hat{V}, \quad (1c)$$

$$\text{div } \hat{E} = 4\pi\rho. \quad (1d)$$

\hat{E} and \hat{H} are the electric and magnetic force vectors, ρ is the charge density while \hat{V} is the velocity vector. $\rho\hat{V}$ is therefore the current density. As usual c stands for the velocity of light in ether. From (1c) and (1d) follows the equation of continuity,

$$\text{div } (\rho\hat{V}) + \dot{\rho} = 0, \quad (2)$$

which means that the charge is conserved. The force per unit charge in the electromagnetic field is given by

$$\hat{F} = \left(\hat{E} + \frac{1}{c} \hat{V} \times \hat{H} \right) \quad (3)$$

where \hat{V} is the velocity of the charge.

The four equations (1) express the quintessence of the classical electromagnetic theory. Faraday's law of the electromagnetic induction is given by (1a). The non-existence of discrete magnetic poles of one kind is assumed by (1b). Ampère's law for the magnetic field due to electric currents and Maxwell's displacement current hypothesis are embodied in (1c) while (1d) stands for the fact that charges are subject to Coulomb's law.

For mathematical convenience the field equations are usually expressed in the following alternative form :

$$\hat{H} \equiv \text{curl } \hat{A}, \quad (4a)$$

$$\frac{1}{c^2} \ddot{\hat{A}} - \nabla^2 \hat{A} = \frac{4\pi}{c} \rho \hat{V}, \quad (4b)$$

$$\frac{1}{c^2} \ddot{\phi} - \nabla^2 \phi = 4\pi\rho, \quad (4c)$$

$$\hat{E} + \frac{1}{c} \dot{\hat{A}} = - \text{grad } \phi, \quad (4d)$$

$$\text{div } \hat{A} + \frac{1}{c} \dot{\phi} = 0. \quad (4e)$$

Although there is a lack of precision about the definitions of \hat{A} and ϕ we introduce the two potentials since it is more convenient to integrate (4b) and (4c).

From the standpoint of the quantum mechanical applications the most important case is that of a moving point-charge e at P for which the retarded potentials,

$$\phi(P, t) = \frac{e}{r + (\hat{V} \cdot \hat{r})/c} \Big|_{t-r/c}, \quad (5)$$

$$\hat{A}(P, t) = \frac{1}{c} \frac{e}{r + (\hat{V} \cdot \hat{r})/c} \Big|_{t-r/c}, \quad (6)$$

were first given by Lienard and Viechart, the right-hand sides being defined for the earlier epoch $t - r/c$. From (4a) and (4d) we get

$$\hat{H} = \frac{1}{r} \hat{E} \times \hat{r} \quad (7a)$$

$$\hat{E} = -e\beta^2/S^3 \left(\hat{r} + \frac{\hat{V}\hat{r}}{c} \right) + \frac{e}{S^3 c^2} \hat{r} \times \left[\left(\hat{r} + \frac{\hat{V}}{c} r \right) \times \hat{V} \right] \quad (7b)$$

where $\beta^2 = 1 - V^2/c^2, \quad S = r + \frac{\hat{V} \cdot \hat{r}}{c}.$

It is obvious from (7a) that the magnetic field is always perpendicular to \hat{r} and \hat{E} . The first part of \hat{E} gives the electrostatic effect which is superseded at great distances from the origin by the second part representing the wave zone. This distinction is valid only if $V/c \ll 1$.

In the case of a harmonic oscillator consisting of a charge e performing small simple harmonic motion the displacement at time t may be taken as

$$x = x_0 \cos Vt \quad (8)$$

\hat{i} being a unit vector. If \hat{j} is the unit vector in the direction of \hat{r} or \vec{PO} where P is an external point and O , the centre of oscillations,

$$\begin{aligned}\hat{H} &= \frac{e}{rc^2} v^3 x_0 \cos vt \sin \theta \frac{\hat{i} \times \hat{j}}{\sin \theta}, \\ \hat{E} &= \frac{e}{rc^2} v^3 x_0 \cos vt \sin \theta \frac{\hat{i} - \hat{j} \cos \theta}{\sin \theta},\end{aligned}\quad (9b)$$

θ being the angle between \hat{i} and \hat{j} .

$$\hat{E} \times \hat{H} = \frac{e^3}{r^2 c^4} v^4 x_0^2 \cos^2 vt \sin^2 \theta (-\hat{j}); \quad (10)$$

$-\hat{j}$ is the direction of the flow of radiation. It is clear that $-\hat{j}, \hat{E}, \hat{H}$ constitute a right-handed system of mutually perpendicular directions. A plane electromagnetic wave can therefore be best described by the equations

$$E_y = a \sin 2\pi\nu (x/c - t) = H_z \quad (11)$$

as $E_x = E_z = H_y = H_x = 0$. The plane of polarization being perpendicular to \hat{E} is xOz .

The energy U and momentum \hat{G}/c of an electromagnetic field are given by

$$U = \frac{1}{8\pi} \int (\hat{E}^2 + \hat{H}^2) d\tau, \quad (12)$$

$$\hat{G} = \frac{1}{4\pi} \int (\hat{E} \times \hat{H}) d\tau, \quad (13)$$

$d\tau$ being the element of volume. In an exclusive field of radiation in which charges do not exist, $|\hat{G}| = U$. This result which follows from Gauss's formula and which makes the energy-momentum vector invariant for the Lorentz transformation is akin to that for a photon viz., $E = pc$, p being the momentum scalar and E the energy.

In the case of the oscillator the energy flow per unit time through an area $r^2 d\omega$ at P becomes

$$\frac{c}{4\pi} \hat{E} \times \hat{H} r^2 d\omega = \frac{d\omega}{4\pi c^3} e^3 v^4 x_0^2 \cos^2 vt \sin^2 \theta. \quad (14)$$

The total energy flow through the spherical surface of radius r per unit time is

$$\frac{1}{4\pi c^3} e^3 v^4 x_0^2 \cos^2 vt \int \sin^2 \theta d\omega = \frac{2}{3c^3} e^3 v^4 x_0^2 \cos^2 vt. \quad (15)$$

Since

$$\frac{\nu}{2\pi} \int_0^{2\pi/\nu} \cos^2 vt dt = \frac{1}{2},$$

the time-average value of the energy flow per unit time is

$$\frac{e^3}{3c^3} v^4 x_0^2. \quad (16)$$

The problems of scattering, emission and absorption are treated by applying classical mechanics to harmonic oscillators and to their interaction with the periodic force of radiation. In the treatment of these and other problems the following defects of the classical theory of radiation are encountered :

(1) The existence of a point charge implies an infinite electrostatic energy. If we take the electron as a sphere of radius r_0 of the order of e^2/mc^2 it has a self-energy and a self-momentum due to its motion in the field as given by the field laws. If the total energy and momentum of the particle and the surrounding field are considered the laws of conservation are fulfilled but the four quantities do not provide a vector to which the Lorentz transformation can be applied. The difficulties of the radius r_0 and of the origin of the electronic mass are interconnected and have not been solved as yet. The physical principles of the old theory as well as the new fail here. The question as to whether mass has electromagnetic origin or not must be examined in the light of the latest results of nuclear physics. Eddington⁵ has derived the equation

$$10x^2 - 136x + 1 = 0 \quad (17)$$

from certain fundamental considerations and he concludes that the ratio of the roots, which is nearly 1845, must be exactly equal to the ratio of the masses of the proton and the electron.

(2) The interaction term for an electron and a light wave contains powers of r_0/λ . As the structure of the electron is neglected in the theory we cannot legitimately discuss an interaction, a large part of which comes from the structure term r_0/λ . Hence we cannot successfully apply the classical theory of radiation if r_0/λ is not small.

These are the intrinsic difficulties of the theory. The phenomenological difficulties have been already mentioned in different contexts. The quantum theory which we describe in the next section shows how light behaves as waves in some situations and as particles in others. Between the dualism of light and that of matter there is one essential difference. Quantization takes us from particles to waves (of probability) in the case of matter and from (electromagnetic) waves to particles in the case of light. If $\hbar = 0$ matter is corpuscular and light is waves.

2. THE QUANTUM THEORY OF RADIATION

It is possible to cast the radiation problem into the Hamiltonian form and then apply the rules of quantization.² Let us first take the simpler case of the pure radiation field for which $\phi = 0$ and \hat{A} satisfies just the two conditions,

$$\nabla^2 \hat{A} - \frac{1}{c^2} \ddot{\hat{A}} = 0, \quad (18a)$$

$$\operatorname{div} \hat{A} = 0. \quad (18b)$$

Consider the radiation enclosed in a cube of side, L , \hat{A} being periodic on the surface. Let L be the unit of length. \hat{A} may be chosen so that

$$\hat{A} = \sum q_{\lambda}(t) \hat{A}_{\lambda}(x, y, z). \quad (19)$$

$$\text{Then if} \quad \dot{q}_{\lambda} + v_{\lambda}^2 q_{\lambda} = 0, \quad (20)$$

$$\nabla^2 \hat{A}_{\lambda} + \frac{v_{\lambda}^2}{c^2} \hat{A}_{\lambda} = 0. \quad (21)$$

The equation (20) is equivalent to that for a harmonic oscillator for which the Hamiltonian is

$$H_{\lambda} = \frac{1}{2} (p_{\lambda}^2 + v_{\lambda}^2 q_{\lambda}^2) \quad (22)$$

p_{λ} being equal to \dot{q}_{λ} . The total energy of the field can be shown to be $\sum H_{\lambda}$.

So far no quantum condition has been used. We have to take the quantum conditions

$$p_{\lambda} q_{\lambda} - q_{\lambda} p_{\lambda} = -\frac{i\hbar}{2\pi}, \quad p_{\lambda} q_{\mu} - q_{\mu} p_{\lambda} = 0, \text{ etc.} \quad (23)$$

In practice there is a change in the procedure because we have to start with a more general expansion of \hat{A} than (19) in order to make the results widely applicable. We now have

$$\hat{A} = \sum (q_{\lambda} \hat{A}_{\lambda} + q_{\lambda}^{\times} \hat{A}_{\lambda}^{\times}) \quad (24)$$

where the cross denotes the conjugate complex character. If

$$Q_{\lambda} = q_{\lambda} + q_{\lambda}^{\times}, \quad P_{\lambda} = -i v_{\lambda} (q_{\lambda} - q_{\lambda}^{\times}) \quad (25)$$

the energy of a single wave is given by

$$H_{\lambda} = \frac{1}{2} (P_{\lambda}^2 + v_{\lambda}^2 Q_{\lambda}^2). \quad (26)$$

The quantum conditions are

$$P_{\lambda} Q_{\lambda} - Q_{\lambda} P_{\lambda} = -i\hbar/2\pi, \quad P_{\lambda} Q_{\mu} - Q_{\mu} P_{\lambda} = 0, \text{ etc.} \quad (27)$$

The eigenvalues of energy are given by

$$E_{\lambda} = (n_{\lambda} + \frac{1}{2}) \hbar v_{\lambda} \quad (28)$$

n_{λ} being a non-negative integer. This produces a difficulty. In a finite volume of radiation there are an infinite number of oscillators and hence the total energy being $\sum E_{\lambda}$ must be infinite even if the quantum numbers n_{λ} are zero. This is impossible. To avoid the difficulty we express (26) as

$$H_{\lambda} = v_{\lambda}^2 (q_{\lambda} q_{\lambda}^{\times} + q_{\lambda}^{\times} q_{\lambda}) = 2 v_{\lambda}^2 q_{\lambda}^{\times} q_{\lambda}. \quad (29)$$

On the classical theory (29) is valid. From (25) and (27) we get

$$H_{\lambda} = 2 v_{\lambda}^2 q_{\lambda}^x q_{\lambda} = \frac{1}{2} (P_{\lambda}^2 + v_{\lambda}^2 Q_{\lambda}^2) - \frac{1}{2} h\nu \quad (30)$$

and hence

$$E_{\lambda} = n_{\lambda} h\nu. \quad (31)$$

The state of the field of radiation is given by the numbers n_{λ} . The functions \hat{A} are of the form

$$\sqrt{(8\pi c^3)} \hat{e}_{\lambda} \cdot e^{i(\hat{X}_{\lambda} \cdot \hat{r})}, \quad (32)$$

where \hat{X}_{λ} gives the direction of propagation and \hat{e}_{λ} is a unit vector which is parallel to the direction of polarization. Since $\text{div } \hat{A}_{\lambda} = 0$, we get

$$(\hat{X}_{\lambda} \cdot \hat{e}_{\lambda}) = 0. \quad (33)$$

Owing to the boundary conditions on \hat{A} , \hat{X}_{λ} takes the values

$$\hat{X}_{\lambda} = \left(\frac{2\pi}{L} n_{\lambda x}, \frac{2\pi}{L} n_{\lambda y}, \frac{2\pi}{L} n_{\lambda z} \right) \quad (34)$$

where $n_{\lambda x}$ etc. are all integers.

$$\frac{v_{\lambda}}{c} = \frac{2\pi}{L} (n_{\lambda x}^2 + n_{\lambda y}^2 + n_{\lambda z}^2)^{1/2}. \quad (35)$$

In the new theory the light quanta appear only through the quantum numbers n_{λ} associated with the harmonic oscillators.

The more complicated problem of the interaction of particles with radiation is also treated by this method. The Hamiltonian is first constructed. It consists of the energy of the particles, the electrostatic field and the wave zone. Only the wave zone can be quantized since the total energy momentum for it can be represented by a four-vector. The electrostatic energy, which is represented by longitudinal waves, does not possess corpuscular properties. A rigorous treatment of the problem has to take account of the four different states in which an electron can exist corresponding to the signs of its energy and spin.

If the question is asked, "What is light, waves or particles?" the modern answer is that it is neither. Light is what it is in the interaction phenomena of matter and radiation. The phenomena are real but all mechanical models devised to explain the phenomena fail. When the frequency of radiation exceeds the value $2 mc^2/h$ an electron pair may be created and it is not therefore right to talk of a pure radiation field of such a high frequency. If one insists upon getting some sort of definition of light the answer may be given that light is an ethereal orchestra of harmonic oscillators conducted by matter according to quantum rules. But such definitions serve no important

purpose in science. Those who know the mathematical theory can do without the definition which is not precise enough. Those who do not understand the theory will not be enlightened by the definition because it is too technical.

The quantum theory of radiation must be looked upon as logically necessary for quantum mechanics to be a complete theory. Consider for example a thought-experiment illustrating the uncertainty principle on which quantum mechanics is based. Suppose that the position of an electron is observed through a γ -ray microscope. The uncertainty of position is of the order of the wave length. If the energy of light were transferred continuously by waves the momentum of the beam which kicks the electron before reaching the eye could be made appreciably small by regulating the incidence of light. In that case the uncertainty relation $\Delta p \Delta q \sim h/2\pi$ would not be valid. The wave theory of light is not therefore consistent with the uncertainty principle.

3. PHENOMENOLOGICAL SUCCESS

We will describe in this section certain phenomena in the explanation of which the quantum theory of radiation has attained a good measure of success. The task will confront us with questions relating to the transition probabilities and hence we consider the general perturbation theory first.

Suppose that the Hamiltonian of a quantum system can be split up into two parts so that

$$H = H_0 + H' \quad (36)$$

where H' is small compared to H_0 and the eigenfunctions ψ_n and eigenvalues E_n corresponding to the latter are known. Thus

$$H_0 \psi_n = E_n \psi_n. \quad (37)$$

We proceed to solve the equation

$$(H_0 + H') \psi = \frac{i\hbar}{2\pi} \frac{\partial \psi}{\partial t} \quad (38)$$

with the assumption that ψ is of the form

$$\psi = \sum b_n(t) \psi_n e^{-2\pi i E_n t / \hbar}. \quad (39)$$

The complication caused by degenerate states will not be discussed here. Substituting for ψ in (38) and using (37) we have

$$\sum \frac{i\hbar}{2\pi} \dot{b}_n \psi_n e^{-2\pi i E_n t / \hbar} = \sum b_n(t) H' \psi_n e^{-2\pi i E_n t / \hbar}. \quad (40)$$

Multiply both sides by ψ_m^* and integrate over the co-ordinate space to get, on using the conditions of orthogonality and normalization,

$$\frac{i\hbar}{2\pi} \dot{b}_m(t) = \sum_n b_n(t) e^{-2\pi i t (E_n - E_m) / \hbar} \int \psi_n^* H' \psi_m d\tau. \quad (41)$$

A rigorous solution of (41) is not an easy matter. On the assumption that H is small and that the system was initially in the unperturbed state given by

$$b_n(0) = 1, \quad b_m(0) = 0, \quad m \neq n,$$

we have

$$b_m(t) = \left[\frac{e^{-2\pi i t (E_n - E_m)/h} - 1}{(E_n - E_m)} \right] H'_{nm}$$

where

$$H_{nm} = \int \psi_m^* H \psi_n d\tau.$$

The probability that the system which was initially in the state ψ_n will be in the state ψ_m at time t is

$$|b_m(t)|^2 = \frac{2 |H'_{nm}|^2}{(E_n - E_m)^2} \left[1 - \cos \frac{2\pi(E_n - E_m)t}{h} \right].$$

Sometimes no direct transition from ψ_n to ψ_m is possible as $H_{nm} = 0$. It is however possible that the two states are connected through intermediate states ψ_p . We have then to find out the probability of the transition from ψ_n to ψ_m by considering all intermediate transitions of the form, $\psi_n \rightarrow \psi_p$ and $\psi_p \rightarrow \psi_m$. For this (41) is split up into two equations:

$$\frac{ih}{2\pi} \dot{b}_p = \sum_{n_0} b_{n_0} e^{-2\pi i t (E_{n_0} - E_p)/h} H'_{n_0 p} \quad (41a)$$

$$\text{and} \quad \frac{ih}{2\pi} \dot{b}_m = \sum_p b_p e^{-2\pi i t (E_p - E_m)/h} H'_{p m}. \quad (41b)$$

When the state ψ_m belongs to a series of continuous states let $\rho_E dE$ denote the number of states with energy between E and $E + dE$. The probability of transition to a state lying in an interval of energy ΔE is, on replacing E_m by E

$$\int_{\Delta E} |b_m|^2 \rho_E dE = \frac{4\pi^2 t}{h} \rho_E |H'_{nm}|^2. \quad (42)$$

Hence the transition probability per unit time is

$$\frac{4\pi^2}{h} \rho_E |H'_{nm}|^2. \quad (43)$$

The problem of natural line breadth has been solved by using this result on transition probabilities and we give therefore an account of it here.

Spectral lines have a natural breadth as otherwise they would not be observable. According to Bohr's theory the breadth is the same for all lines. This disagrees with observation. The new theory provides an explanation of the varying natural breadth of lines. The classical treatment attributes this phenomenon to the electronic self-force. A calculation of the self-force for an oscillating electron gives the more rigorous equation,

$$\ddot{x} = -\nu_0^2 x - \Gamma \dot{x} \quad (44)$$

where $\gamma = \frac{2}{3} e^3 v_0^2 / mc^3$. (45)

If we assume that

$$\gamma \ll \nu_0 \quad (46)$$

the solution of (44) may be written down as

$$x = x_0 e^{-\gamma t/2} \cos \nu_0 t. \quad (47)$$

An approximate estimate of the energy is

$$U = \frac{1}{2} m (\dot{x}^2 + \nu_0^2 x^2) = \frac{1}{2} m \nu_0^2 x_0^2 e^{-\gamma t} = U_0 e^{-\gamma t}. \quad (48)$$

By convention $1/\gamma$ is taken as the average lifetime of an oscillator. In this time the energy diminishes in the ratio $e : 1$. Moreover, as

$$\int_0^\infty U_0 e^{-\gamma t} dt = U_0 \cdot 1/\gamma$$

the total action of the oscillator is equivalent to the action of a simple harmonic oscillator of energy U_0 and lifetime $1/\gamma$. The electrical force vector corresponding to (9b) is the real part of

$$\hat{E} = \hat{E}_0 e^{-\gamma t/2} \cdot e^{i\nu_0 t} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{E}(\nu) e^{i\nu t} d\nu$$

where

$$\begin{aligned} \hat{E}(\nu) &= \frac{1}{2\pi} \hat{E}_0 \int_0^\infty e^{-\gamma t/2} e^{i(\nu_0 - \nu)t} dt \\ &= \frac{\hat{E}_0}{2\pi} \frac{1}{i(\nu_0 - \nu) - \gamma/2}. \end{aligned}$$

The intensity distribution is given by

$$I(\nu) \simeq |\hat{E}(\nu)|^2 = I_0 \gamma / 2\pi \frac{1}{(\nu - \nu_0)^2 + \gamma^2/4} \quad (49)$$

so that

$$\int_{-\infty}^{\infty} I(\nu) d\nu = I_0. \quad (50)$$

Over the range $\nu - \nu_0 = \pm \gamma/2$ the intensity is not less than half the maximum value. Hence γ is known as the half-maximum breadth. Since $\Delta\nu_0 = \gamma$ by (45)

$$\Delta\lambda = \frac{c}{\nu_0^2} \Delta\nu_0 = \frac{2}{3} \frac{e^2}{mc^2} = 1.7 \times 10^{-4} \text{ \AA}. \quad (51)$$

Thus the broadening in wave length is found to be of the order of 10^{-13} cm. and independent of the wave length.

The relation between the lifetime and the half-maximum breadth follows from the uncertainty relation,

$$\Delta E \Delta t = h/2\pi, \quad (52)$$

of quantum mechanics. If Δt is the average lifetime of a state only so much time is available for finding its energy. Hence the uncertainty in the energy determined is

$$\Delta E = \frac{h}{2\pi} \cdot \frac{1}{\Delta t} = \frac{h\gamma}{2\pi}. \quad (53)$$

To determine γ we proceed with a Hamiltonian for the joint system of atoms and radiation, arrive at the equation (41) and then, from the initial conditions get the form of $b_n(t)$. Let $b_n(0) = 1$ and $b_n(t_1) = 0$. Then the average lifetime is

$$\frac{1}{\gamma} = \int_0^{t_1} -t db_n. \quad (54)$$

The lifetime obviously depends upon the transition probabilities and the exact relation between the lifetime of a state ψ_n and the probabilities per unit time W_{nm} for transitions from ψ_n to lower levels ψ_m is

$$\gamma = \sum_m W_{nm}. \quad (55)$$

This gives us the thickness associated with each level,

$$\Delta\nu_n = \frac{1}{2\pi} \sum_m W_{nm}. \quad (56)$$

The line breadth of an emitted line, $\psi_n \rightarrow \psi_{n'}$ is

$$\Delta\nu_n + \Delta\nu_{n'}. \quad (57)$$

It is possible on the new theory even for a weak line to be broad. Moreover, the breadth varies from line to line. The shape of the absorption line is the same as that of the emission line if there is no appreciable variation in the intensity of the incident line in the region of the line breadth. It may be mentioned in passing that several other factors exist which considerably influence the shape of the line³; for instance, the Doppler effect and collisions with other atoms and molecules.

An excited atom, left to itself, may emit radiation and attain a state of lower energy. Such spontaneous emission could not be explained by the correspondence principle and the classical theory. Induced emission in the field of atoms and radiation could be explained without any reference to the quantum theory of radiation as Einstein first did. The problem of spontaneous emission is solved along with the other problems of the interaction of matter and radiation by constructing a quantum Hamiltonian for the joint system of matter and radiation. Dirac's relativistic treatment of the electron gives for the joint system

$$H = H_0 + \hat{H}$$

$$\text{where } H_0 = \sum_k \left[(\hat{\alpha}_k, \hat{P}_k - e_k \hat{A}) + \beta_k \mu_k + e_k \phi \right] + \frac{1}{2} \sum_i \sum_j \frac{e_i e_j}{r_{ij}} + \sum_\lambda n_\lambda \quad (58)$$

$$\hat{H} = - \sum_k e_k (\hat{\alpha}_k, \hat{A}(k)).$$

\hat{A} represents the vector potential of the external field and the radiation, ϕ the scalar potential of the external field and $\hat{\alpha}_k, \hat{\beta}_k$ the Dirac matrices for the k^{th} particle. The pure radiation field is described by the term, $\sum_{\lambda} n_{\lambda} \frac{k_{\lambda}}{\lambda}$.

The probability of emission is given by the formula (43) of which the term $|H_{nm}|^2$ can be shown to have a factor $1 + n_{\nu}$ where the suffix is

$$\nu = \frac{E_n - E_m}{h}.$$

Hence even if $n_{\nu} = 0$, that is, when there is no radiation of frequency ν present in the field emission of that frequency can occur. This is the explanation of spontaneous emission.

Recent years have brought to light some curious phenomena of γ -rays which are typical quantum effects. The first is that of multiple processes. Suppose that a light quantum whose momentum is \hat{K}_0 is scattered by an electron at rest with energy $\mu = mc^2$ and that the scattered radiation consists of two quanta with momenta \hat{K}_1 and \hat{K}_2 . If \hat{P} is the momentum imparted to the electron and E is its energy after the impact, the laws of conservation give

$$\begin{aligned}\hat{K}_0 &= \hat{K}_1 + \hat{K}_2 + \hat{P}, \\ K_0 + \mu &= K_1 + K_2 + E.\end{aligned}$$

The final stage denoted by $F(\hat{K}_1, \hat{K}_2, \hat{P})$ can be reached through six different intermediate stages. The stages may be described in pairs as follows:

(1) \hat{K}_0 absorbed. The electron has a momentum $\hat{P}' = \hat{K}_0$ and (2) \hat{K}_1 is emitted. The electron has a momentum $\hat{P}'' = \hat{K}_0 - \hat{K}_1$. When \hat{K}_2 is emitted the final stage is reached. But we could also have as intermediate stages

- (1) \hat{K}_0 absorbed with electron momentum \hat{P}' ,
 and (3) \hat{K}_2 emitted with electron momentum $\hat{P}'' = \hat{K}_0 - \hat{K}_2$,
 or (4) \hat{K}_1 emitted with the electron momentum $-\hat{K}_1$,
 and (5) \hat{K}_2 emitted with the electron momentum $-\hat{K}_1 - \hat{K}_2$,
 or (6) \hat{K}_2 emitted with the electron momentum $-\hat{K}_2$.

The transition probabilities have to be calculated taking note of all these transitions. For $K_0 \ll \mu$ the double process is extremely rare. For $K_0 \gg \mu$ the probability of this process is $\frac{1}{137}$ of that of the single process. Processes are also possible scattering one impinging photon into n photons but their probabilities are extremely small. One witnesses multiple processes in the interaction of cosmic rays with matter.

The creation of electron pairs and also the annihilation which may sometimes give rise to Bremsstrahlung (continuous X-ray emission) are features of the interaction of hard γ -rays with matter. When a γ -ray quantum with energy exceeding $2mc^2$ interacts with an electron in a negative-energy state the latter may rise to a state of positive energy. In this process both the energy and the momentum are conserved only if a third particle participates in the interaction. The reverse process, that of annihilation, takes place when an electron with positive energy has momentum equal and opposite to that of a positron. The energy is released in the form of two quanta with equal and opposite momenta. The annihilation may lead to one quantum of radiation if the electron with positive energy is bound to a nucleus and it may also happen that the energy released in annihilation is absorbed in an atom. When hard cosmic rays interact with the atmospheric matter electron pairs are created by pairs of scattered photons with combined energy exceeding $2mc^2$ which increases the scattering of the rays. We have here the phenomenon of light scattered by light about which we will have to say something in the next section. A fast electron moving in the field of an atom gets deflected and the energy lost is emitted as Bremsstrahlung. The latter may consist of one or more quanta. The calculation of transition probabilities of these processes goes far beyond the scope of the older theories of radiation and the correspondence principle.

Cosmic-ray studies and nuclear physics have revealed phenomena which bear out the new theory of light, the hole theory of the positron and the present treatment of the interaction of matter and radiation. But our refined optical instruments have also brought us face to face with situations which question the soundness of the logic of the new theory. In the following section, we will deal with the difficulties which the theorist has to tackle in this field.

4. PECULIAR FEATURES AND DIFFICULTIES

The difficulties that are inherent in the nature of an electron persist even in the quantum theory. Although the electron is treated as a point charge nothing is known about the laws that hold good within a distance of 10^{-13} cm. of the charge. It is not proper to treat the electron as a mathematically abstract point within a micro-microscopic neighbourhood of itself although that would be justifiable in problems where inter-electronic distances of the order of 1\AA are involved. A certain sphere of influence of radius of the order of 10^{-13} cm. is usually identified with the electron. The quantum laws do not hold good here. In fact they fail just outside this region also. So the problem of electronic mass and self-energy still remain unexplained. But the matter does not end here. The difficulties are aggravated near the edge of the electron by those very considerations which are responsible for the success of the quantum theory at greater distances. The interaction energy of

matter and radiation \hat{H} has the average value zero to the first approximation. But if the second order terms are retained in the interaction of an electron with radiation and the average value is determined it is found to be

$$W = \frac{\hbar^2}{137\pi mc^3} \int_0^\infty \nu d\nu = \frac{2\hbar e^2}{mc^3} \int_0^\infty \nu d\nu. \quad (59)$$

This is known as the transverse self-energy of the electron as it is due to the interaction with light waves. Why this term is called self-energy should be made clear because it represents a typical quantum effect.

The field of radiation is characterised at each point by an observable N whose eigenvalues are the quantum numbers for the orchestra of fictitious oscillators. N does not commute with the electric force vector \hat{E} . Hence the uncertainty relation

$$\Delta N \Delta \hat{E} \sim \hat{E}$$

holds good. Suppose that there are no photons in the field so that $N = 0$ and $\hat{E} = 0$. But the uncertainty relation says that even in such a case there may be fluctuations $\Delta \hat{E}$ such that the average value of \hat{E} is zero. The interaction of the electron with these fluctuations is responsible for the transverse self-energy when the field of radiation is absent. When $\hbar = 0$, $W = 0$ and hence the origin of W is to be traced to the quantum hypothesis.

The second approximation fails for precisely the same kind of physical reasons for which the classical theory breaks down when the condition

$$\lambda \gg r_0 = e^2/mc^2 \quad (60)$$

is violated. One interpretation of this condition is already given. It has to be satisfied for the validity for the expression of the self-force used in (44). Further (46) is practically equivalent to (60) since the actual statement of the former reduces to

$$\lambda \gg 2/3 r_0. \quad (61)$$

In the form in which (46) is stated the condition means that the life of an oscillator must be large compared to the period of oscillation as otherwise it would be meaningless to talk about an oscillator. For similar reasons the condition $t \gg \hbar/E_0$ has to be fulfilled in connection with (42) if the existence of stationary states is to have a meaning. \hbar/E_0 gives the period of revolution of the electron.

The interaction with atoms of electrons with energies of the order of $100mc^2$ or more is inferred from cosmic-ray photographs. The classical theory breaks down for electrons of energies exceeding $137mc^2$ as the condition (60) is no longer valid. The quantum theory gives an exceedingly higher loss of energy by radiation for the interaction of electrons having $E > 150mc^2$. The theory is in good agreement with experimental results for $E < 150mc^2$.

As we have already remarked above electron pairs are created in a field of radiation with

$$\lambda < \lambda_0 = h/2mc. \quad (62)$$

This limit on quantum electrodynamics *in vacuo* is much more stringent than the limit (60) on the classical theory. What is most interesting about this limit is that it vanishes if we put $\hbar = 0$ or take the velocity of light as infinite. It is interesting to note that the limit λ_0 is independent of the electronic charge. Incidentally, consider a wave packet of light or of a single electron; when it is of an extension smaller than λ_0 its Fourier expansion will involve waves corresponding to states of energy exceeding $2mc^2$; that implies the existence of an indefinite number of electron pairs. It is hazardous therefore to attempt to draw conclusions from wave packets of extension less than λ_0 .

The properties of the positive electron and its interaction with radiation are not yet clearly understood. We have already discussed elsewhere the theory of the positron and we will content ourselves merely with the remark that the hole theory gives correct results either where first order effects are concerned or when the interacting field is very weak.

A light quantum has an infinite self-energy. This is due to the fact that its identification with itself can be accomplished through intermediate stages of multiple processes involving the creation of an infinite number of pairs. The hole theory itself can in general be applied only when a finite number of pairs is created and the probability of creation is very small.

After a critical survey of these difficulties Heitler² concludes: "The present quantum electrodynamics gives correct results if applied only to the first approximation in the interaction of light with both a single free electron and with the infinite sea of electrons in negative energy states. Nevertheless, we cannot in general hope to obtain correct results for light in which wave lengths $\lambda \sim r_0$ occur".

Further research is necessary before we understand how the electronic radius r_0 or e^2/mc^2 is independent of \hbar . In the mystery of this apparent independence of r_0 and \hbar probably lies the root of future theoretical developments. It is gratifying to note that in spite of many revolutions in thought the principles of momentum and energy are still applied to individual processes. One naturally hopes to count on them in the spade-work that lies ahead in this field of phenomena.

5. THE POSITIVE ELECTRON

On theoretical grounds the existence of the positron or the positive electron was anticipated two or three years before it was actually discovered by Anderson in a cosmic-ray photograph.⁴ The relativistic formula for the energy of an electron in free space is

$$E = \pm \sqrt{(\hat{p}^2 + m^2 c^4)}$$

in the usual notation. Negative values of energy have therefore to be admitted in the relativistic treatment of an electron. The possible values of E are E_+ or E_- such that

$$E_+ \geq mc^2, \quad E_- \leq -mc^2.$$

The probability of a transition from E_+ to E_- is finite. Actually the transition may occur with an emission of radiation or as a result of collision. It is necessary, therefore, to understand clearly the physical significance of electrons in negative energy states.

The most obvious interpretation of an electron in the state E_- suggests properties which have never been observed of electrons. For instance an electron with energy E_- and another with energy E_+ would repel each other according to the classical theory. But the force of repulsion on the former would induce in it an acceleration as if it were subject to an attraction and in the latter, the usual effect. The mischief is caused by the apparent change in the sign of the mass of the electron with negative energy. It is only when the mass is positive that the acceleration is in the same sense as the external force. The interpretation of the state E_- on Dirac's hole theory avoids the difficulty caused by a negative mass. But the particles whose properties were discussed on this theory were unknown in 1930 and Dirac's first suggestion that they might be protons could not be seriously entertained. Although Dirac's wave equation giving the magnetic moment of the electron appeared mathematically satisfactory in every way there still remained the unsatisfactory feature of negative energy states which was finally removed by the hole theory when the positron was discovered. Of course, the latter theory has its own logical difficulties today. The following are the two assumptions on which it is based :

(1) Ordinarily, that is, in the absence of an external field all negative energy states are occupied. The electrons occupying the states partly constitute, so to say, the substratum of the universe in which the electrons move. Hence, normally, by the Pauli principle, it is not possible for an electron to jump to a level of negative energy.

(2) The field laws of classical electrodynamics and mechanics are supposed to represent the full effect of the substratum inasmuch as the electrons occupying the negative energy levels do not disturb the classical calculations of the force, energy and momentum. If, however, an electron leaves one of the so-occupied levels and climbs to a state of positive energy, the resulting deficiency or hole in the substratum reveals itself in respect of charge, momentum and energy.

Suppose that an electron with energy E_- and momentum \hat{p} leaves one of the occupied levels. The absence of the electron must make itself felt as a deficiency in the effect of the substratum. By the second assumption this deficiency or hole in the occupied zone gives to the external field a charge e

if $-e$ is that of the electron, a momentum $-\hat{p}$ and an energy $|E_-|$. The hole is therefore interpreted as a positive electron with charge e , momentum $-\hat{p}$ and energy $|E_-|$. The creation of the hole or the positive electron is obviously accompanied by the appearance of a new electron, the one that had originally filled the hole, in one of the states of positive energy. An electron pair thus comes into existence simultaneously. Similarly when an electron acquires charge, momentum and energy, exactly what is required to fill the hole, the electron pair may be annihilated.

The arguments regarding the creation or annihilation are valid only if a finite number of electron pairs are considered. The theory of the positron is not valid when an infinite number of such pairs are supposed to be featuring in any phenomenon. For the very laws of mechanics and electrodynamics on which the arguments depend break down when the substratum is appreciably disturbed by the existence of an infinite number of holes. In the first place it is not logical to assume that an external field of the classical type can disturb the substratum itself to the extent of producing a hole in it. A radiation of energy exceeding $2mc^2$ is needed for disturbing the substratum. The theory of radiations of this order of energy, that is, of wave lengths less than $\hbar/2mc$ cannot be rigorously treated on classical assumptions because the latter are intimately related to the possibility of all negative energy states being occupied. In other words, the two assumptions of the hole theory serve the purpose of only reconciling the negative energy states with classical mechanics and electrodynamics; and the supposed mechanism of the creation and annihilation shows that the assumptions are plausible to a limited extent beyond which range phenomena transcending the classical theory. The difficulty of E_- states is closely bound up with those of radiation of high energies. Until they are solved one cannot determine the interaction between a strong field and the positrons. For weak fields the results of interaction obtained from the classical theory are true, at least, qualitatively. One such elementary qualitative result which was responsible for the discovery of the positron is that positive and negative electrons develop opposite curvatures in an external magnetic path.

The first serious break in the classical theory appeared when Dirac formulated the laws of quantum mechanics in the old symbolic form of Poisson brackets but with the refinement of non-commutative operations, heralding a new philosophy of the states and observables of systems in harmony with experimental limitations. A further advance was achieved when the relativistic wave equation of the electron was cast into a linear form so that four distinct states became possible by virtue of the different signs of the spin and energy. It is significant that Dirac's relativistic equation can be expressed as equivalent in form to the Maxwell equations of electrodynamics. Because, from this stage onwards, a joint treatment becomes necessary of matter and radiation. The first important phase of the theory of radiation concludes

with the second quantization which, in substance, has been described by the conditions (27). The difficulties of high energy in this theory and those of negative energy states in the other can be satisfactorily removed only in a unified treatment of matter and radiation.

The developments outlined by us have been marked by an unshaken faith in the classical principles of conservation for matter and radiation. Whenever we are concerned with a partial knowledge of both kinematical and dynamical features of a system wave mechanics answer questions in terms of probability and the phenomena are governed by statistical laws. But deterministic laws regarding the energy and momentum of a single particle still hold good provided no space-time description is contemplated. The concepts of the neutron and the neutrino and our explanation of various nuclear processes are characterized by the success of the conservation principles for particles. If the principles fail, as Bohr once suggested in connection with the uneven β -rays, it will only mean that we cannot isolate the particle in question. In the future theoretical developments of the micro-microscopic phenomena conservation principles for isolated particles are bound to be replaced by field laws. Dimensionless numbers such as $hc/2\pi e^2$, which Eddington has tried to explain, may possibly be understood in fundamental terms only when the developments have occurred.

REFERENCES

1. CONDON AND SHORTLEY, "*The Theory of Atomic Spectra*" pp. 79-110 (1935).
 2. HEITLER, W., "*The Quantum Theory of Radiation*", articles 6, 12, 19, 24 and p. 236 (1936).
 3. MARGENAN AND WATSON, "*Rev. Mod. Phys.*", 8, 1, pp. 22-29 (1936).
 4. MILLIKAN, R. A. "*Electrons (+ and -) etc.*", pp. 320-59 (1935).
 5. EDDINGTON, A. S., "*Relativity Theory of Protons and Electrons*" p. 300 (1936).
-

IX QUANTUM THEORIES OF FORCES

1. The Nature of Chemical Forces. 2. Collectivization of Electrons. 3. Some Typical Methods of Calculation. 4. Other Forces of Quantum-Mechanical Nature.

1. THE NATURE OF CHEMICAL FORCES

The concept of valence¹ is introduced to account for the whole number relationship of chemical combinations. It has been most successfully used in organic chemistry. The forces of valence are essentially electrical and hence a wave mechanical theory of electrons in atoms must be competent to explain quantitatively and qualitatively various questions regarding the formations of molecules. One simple criterion decides as to whether a given number of atoms can combine to form a stable molecule or not. If the total energy of the molecule is less than the sum of the free-space energies of the constituent atoms a stable molecule is formed and the energy that is released as the molecule comes into existence is called the heat of formation or molecular heat. The criterion furnishes a very simple explanation of complicated processes such as polymerization, oxidation and condensation and of unusual compounds like BeH. A treatment of the n -body problem can answer therefore all questions regarding valence including anomalous and directed valence, homopolar bonds and activation energies. But the rigorous treatment of the n -body problem has difficulties practically insurmountable and one has to resort to crude approximations. Hence the results are of a qualitative nature only. The task of approximation is considerably lightened by the fact that the electrons of the closed inner shells and subshells can be regarded as virtually absent if the method of the self-consistent field is adopted. When quantitative results are desired the inner electrons prove to be a nightmare.

On this view all valence problems are really problems of adjustment of a joint family of atoms. Usually, for any molecule, the adjustment has two principal features. Firstly, the nuclei of participating atoms assume some definite relative positions. Thus the hydrogen nuclei of H_2O are at equal distances from the oxygen nucleus where an angle about 106° is subtended. Similarly the spatial arrangement of the nuclei of CH_4 presents a tetrahedral character. Secondly, the electrons in the outermost open orbits of the atom which are called free or valence electrons are redistributed about the different nuclei in accordance with the Fermi-Dirac statistics. The effect

of this is that some of the valence electrons are elevated to higher energy levels, some go to lower levels while the others remain more or less undisturbed. The electrons that attain a higher level do not favour the formation of a molecule and are therefore called anti-bonding. Those that occupy levels of lower energy are known as bonding electrons. The rest are non-bonding. The inner electrons are generally non-bonding but there are exceptions. In the case of copper, silver and gold one or two inner electrons become bonding and hence we have all of them divalent and the third also trivalent. There is no longer any meaning in the abstract to the question, "What is the valency of iron or copper?" For the valency, that is, the number of bonds depends upon the other atom or atoms which participate in a chemical combination with it. This is the reason why oxides and hydrides reveal different valencies of elements. Herzberg has given the following semi-empirical rule for valency bonds :

No. of bonds = $\frac{1}{2}$ (No. of bonding electrons—No. of anti-bonding electrons.)

(1) The rule has been verified in several cases. In the case of nitrogen there are ten outer electrons for two atoms; two are anti-bonding and eight bonding. Thus (1) explains the triple bond of nitrogen, $N \equiv N$.

If E is the energy of an atom with n electrons it is generally supposed to be expressible in the form

$$E = E_0 + \sum_{r=1}^n E_r \quad (2)$$

where E_0 is the energy of the nucleus and E_r , that of the r^{th} electron. This assumption cannot be rigorously true as the interaction energy is distributed over the terms of the right-hand side in (2). If the n^{th} electron is removed the energy of the ionized atom is not just $(E - E_n)$; nor does the energy of the r^{th} electron remain E_r . But it may be possible to express the new energy as

$$E' = E'_0 + \sum_{r=1}^{n-1} E'_r, \quad (3)$$

where all the symbols are new as they represent a redistribution of the interaction energy. The energy of an electron in the atom is thus a fictitious concept introduced for the sake of the simplicity and convenience of the formula (2). A rigorous determination of E is impracticable and hence one has to be content with an approximation. The self-consistent field method is usually very useful for calculating E_r when the interaction energy is negligible. In view of this it is difficult to decide whether an electron is bonding, non-bonding or otherwise especially when the number of valence electrons is large.

The spin serves as a valuable index to the nature of the valence electron. The part played by the electron in chemical combination was very much misunderstood in the earlier quantum theories of valence. The bonds were

attributed to the magnetic forces of the spins. But this could not be right as these forces are comparatively very small. According to the earlier theory of heteropolar valency as given by Lewis for dissimilar atoms each nucleus tends to be surrounded by complete shells of even members of electrons as in an inert gas atom. It was interpreted later as a tendency of the electrons with opposite spins to form pairs. It was also supposed to explain the electron-pair bonds of homopolar valency. If the matter were to be decided merely by the pairing of bachelor electrons of opposite spins both BH_3 and B_2H_6 should exist. We know, however, that the former is not normally possible. O_2 is a stable molecule in which all the electrons are not paired and hence there is a resultant magnetic moment. As against the theory of the electron-pair bonds there is the example of H_2 in which only one electron is responsible for the chemical bond. Quantum mechanics has shown that the electron spin controls the phenomena of chemical combination through the Pauli exclusion principle. For the choice of the ψ -function upon which the magnitude of the interaction depends is apparently influenced by the spin. Heitler and London's method shows why the valency in many cases is equal to the number of bachelor electrons in the outer shell. The secret of the importance of the spin lies in the fact that the exchange energy and the spin interaction can be simultaneously diagonalized.

One expects that the number of bachelor electrons in the outer shell of an atom to be zero or one. Considerations of energy suggest that if there are even valence electrons they would be all paired if the energy is to be a minimum. Similarly, when there are odd electrons it is expected that only one of them would be a bachelor electron. The exception is fulfilled in the case of lighter atoms. But even in their case it is found that in the adjustment preceding the formation of a molecule some of the electron pairs may be divorced. The various possible valencies exhibited by chlorine, which are 1, 3, 5 and 7, are explained by this process. Carbon is known to be quadrivalent. In the ground state there are only two bachelor electrons in the $2p$ shell. In the presence of another atom one pair is usually divorced so that four electrons with parallel spins appear in the carbon atom.

When two electrons are very close to each other their average interaction energy is large and the formula (2) fails. Heavy elements with a large number of valence electrons illustrate the situation. If the Fermi-Dirac statistics is to be obeyed and if the building-up principle is not to be violated two such electrons must move in different spatial orbits. Their spins may or may not be parallel. But under the influence of an external electrical field the spins may be made parallel. That is how the large number of valencies of transition metals, iron and others are accounted for.

Various quantum mechanical theories have been proposed to remove the difficulties of the n -body problem which vitiate the theoretical conclusions

regarding chemical combinations. There are the directional and non-directional theories of Dirac and Slater; the theory of molecular orbitals first used by Lennard-Jones and that of Heitler and London are also at the disposal of the research worker. We shall give an account of the latter theories with reference to the ionized molecule H_2^+ .

We have already referred to bonding, non-bonding and anti-bonding electrons. The theory of H_2^+ is essentially one of a 'vagabonding' electron. The first account of the vagabonding electron appeared in Heisenberg's papers but now it is recognized to be definitely wrong. Heisenberg's treatment⁸ of the problem was based on the analogy of two simple pendulums supported by a horizontal string.⁹ If one of the pendulums is set in motion its motion is gradually transferred to the other so that after some time the first is brought to rest. The motion is thus completely transferred from one pendulum to the other at regular intervals. Heisenberg argues that the electron of H_2^+ similarly changes its nucleus with a frequency ν . The wave function for the electron was shown to be

$$\psi = \frac{1}{2} \left[\psi_A^0 \left(e^{-2\pi i W_1 t/h} + e^{-2\pi i W_2 t/h} \right) + \psi_B^0 \left(e^{-2\pi i W_1 t/h} - e^{-2\pi i W_2 t/h} \right) \right] \quad (4)$$

so that the probability density becomes

$$\psi^* \psi = |\psi_A^0|^2 \cos^2 2\pi \nu t + |\psi_B^0|^2 \sin^2 2\pi \nu t, \quad (5)$$

where $2h\nu = W_1 - W_2$. ψ_A^0 and ψ_B^0 are certain functions of spatial position. The coordinates of position measured relative to the nucleus A appear in ψ_A^0 only and similarly ψ_B^0 is defined for the other nucleus. At $t = 0$, the electron is about the nucleus A and at $t = 1/\nu$ it goes over to the field of the other. W_1 and W_2 are the energies corresponding to the symmetrical and anti-symmetrical wave functions. This explanation is rejected on the ground that the formation of H_2^+ is a static phenomenon and that by the principles of wave mechanics transitions between symmetrical and anti-symmetrical states cannot happen. Hence it is wrong to assume the wave function (4) for the valence electron.

Frenkel has used the word 'collectivization' to describe the relation of the valence electron to the two nuclei. Every chemical combination is a collectivization of electrons by the nuclei which also adjust their relative positions so that the total energy becomes a minimum. This idea of sharing up electrons is quite old: it appeared in Lewis's theory. But only wave mechanics has worked it out successfully. We are now going to see how, in particular, one electron can be shared up between the two nuclei.

2. COLLECTIVIZATION OF ELECTRONS

In the method^a of molecular orbitals we start with the assumption that the wave function of an atom with n electrons can be expressed as

$$\psi = \psi_1(x_1, y_1, z_1) \psi_2(x_2, y_2, z_2) \cdots \psi_n(x_n, y_n, z_n) \quad (6)$$

where $\psi_r(x_r, y_r, z_r)$ is the molecular orbital for the r^{th} electron. It is assumed that ψ_r is a solution when other electrons are absent. As a particular case consider two hydrogen atoms with two valence electrons. If we call the atoms A and B we may take

$$\psi_1 = a \psi_A(x_1, y_1, z_1) + b \psi_B(x_1, y_1, z_1), \quad (7)$$

ψ_A and ψ_B being possible wave functions for an electron in the fields of the nuclei of A and B respectively. If ψ_A and ψ_B are real and normalized

$$\int \psi_A^2 d\tau = \int \psi_B^2 d\tau = 1, \quad (8)$$

and if ψ_1 is also normalized

$$a^2 + b^2 + 2ab I = 1 \quad (9)$$

where

$$I = \int \psi_A \psi_B d\tau. \quad (10)$$

If the other electron is in an equivalent state

$$\psi_2 = a \psi_A(x_2, y_2, z_2) + b \psi_B(x_2, y_2, z_2), \quad (7')$$

$$\text{and } \psi = a^2 \psi_A(1) \psi_A(2) + b^2 \psi_B(1) \psi_B(2) + ab \left(\psi_A(1) \psi_B(2) + \psi_A(2) \psi_B(1) \right). \quad (11)$$

In the method of collectivization which is developed by Lennard-Jones, Mullikan and others only the first two terms on the right-hand side of (11) are considered as significant. These terms are called ionic since the electrons are both in the same atom A or B . It is expected however that the electrons would avoid each other on account of mutual repulsion. Hence in taking only the ionic terms as is done generally in this method one commits the mistake of ignoring the repulsion between the electrons. In the Heitler and London method this mistake is avoided by rejecting the first two terms and retaining only the coefficient of ab in (11). If the distance between the atoms is large the rejection of ionic terms provides a good approximation but when the distances are small the second method breaks down.

Consider two hydrogen nuclei fixed at points $(\pm \frac{1}{2} R, 0, 0)$, R being the distance between them. For an electron in the field of the nuclei let a possible wave function be

$$\psi(x, y, z) e^{-2\pi i E t / \hbar}. \quad (12)$$

As the probability density $|\psi|^2$ must be symmetrical with respect to the protons

$$\psi(x, y, z) = \pm \psi(-x, y, z). \quad (13)$$

If there is only one electron in the field we have, following (7)

$$\psi(x, y, z) = a\psi_0(\hat{r}_A) + b\psi_0(\hat{r}_B) = a\psi_A(x, y, z) + b\psi_B(x, y, z) \quad (14)$$

where the vectors \hat{r}_A and \hat{r}_B , measured from the two nuclei A, B to the probable position of the electron can be expressed in terms of x, y, z . By (13) it is clear that

$$a = \pm b. \quad (15)$$

If ψ_0 is normalized and

$$I = \int \psi_0(\hat{r}_A) \psi_0(\hat{r}_B) d\tau, \quad (16')$$

we may choose for ψ either of the normalized wave functions,

$$\psi_s = \frac{\psi_0(\hat{r}_A) + \psi_0(\hat{r}_B)}{\{2(1 + I)\}^{1/2}} \quad (16)$$

$$\text{or} \quad \psi_a = \frac{\psi_0(\hat{r}_A) - \psi_0(\hat{r}_B)}{\{2(1 - I)\}^{1/2}}. \quad (16'')$$

The wave equation which ψ has to satisfy is

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2} (E - V) \psi = 0 \quad (17)$$

where

$$V = -\frac{e^2}{r_A} - \frac{e^2}{r_B}. \quad (18)$$

Since $\psi_0(\hat{r}_A)$ satisfies

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2} \left(E + \frac{e^2}{r_A} \right) \psi = 0 \quad (19)$$

and $\psi_0(\hat{r}_B)$ a similar equation it follows that (14) is not a rigorous solution of (17). The condition (15) improves the degree of approximation of ψ . Moreover an exact solution of (17) is too difficult and too complicated for application.

So far nothing has been said about the choice of the solution of (19). Let us take the wave function ψ_{100} corresponding to the ground state of the hydrogen atom.

$$\psi_0(\hat{r}_A) \equiv \psi_{100} = (\alpha^3/\pi)^{1/2} e^{-\alpha r_A} \quad (20)$$

where

$$\alpha = 4\pi^2 m e^2 / h^2. \quad (21)$$

Using polar coordinates with reference to the midpoint of the nuclei and the line AB, we have

$$\psi_0(\hat{r}_A) = (\alpha^3/\pi)^{1/2} e^{-\alpha(r + \frac{1}{2} R \cos \theta)} \cong (\alpha^3/\pi)^{1/2} e^{-\alpha r} (1 - \frac{1}{2} R \alpha \cos \theta). \quad (22)$$

It is assumed that R is very small compared to r . So

$$\psi_A(x, y, z) = (\alpha^3/\pi)^{1/2} e^{-\alpha r} (1 - \frac{1}{2} R \alpha \cos \theta). \quad (23)$$

Similarly we get

$$\psi_0(\hat{r}_B) \cong (\alpha^3/\pi)^{1/2} e^{-\alpha r} (1 + \frac{1}{2} R \alpha \cos \theta) \quad (24)$$

and that gives $\psi_B(x, y, z)$ also. The integral (10') can now be evaluated.

$$\begin{aligned} I &= \frac{\alpha^3}{\pi} \int e^{-2\alpha r} \left(1 - \frac{1}{4} \alpha^2 R^2 \cos^2 \theta \right) d\tau \\ &= 1 - \frac{\alpha^2 R^2}{12}. \end{aligned} \quad (25)$$

The very form of the wave function (12) is that of a static distribution of probability. Since the energy E of the electron is fixed the uncertainty principle makes the time indefinite. According to the description of phenomena by ψ there is, in the present case, no such thing as an electron having one position or one probability distribution at one time and another at a different instant. We get a static probability density $|\psi|^2$ which gives rise to the conventional interpretation providing the charge density of the electron cloud in the form $e|\psi|^2$. This is merely a pictorial way of describing a purely mathematical situation in familiar terms. It has been already remarked that Schrödinger's first interpretation of ψ failed because the charge cloud was supposed to possess the properties of an actual charge distribution as in electrostatics. A warning being given regarding the current use of the term 'charge cloud' for the expression $e|\psi|^2$ we will now show how it helps us to understand the chemical forces.

When R is very large I becomes negligible and the functions (16) and (16') give probability distributions such that there are two practically distinct charge clouds each about a nucleus. Here is one of the simplest cases of collectivization, one electron being shared half and half by two atoms. The cloud about A has the density $\frac{1}{2} e |\psi_0(\hat{r}_A)|^2$, that about B having $\frac{1}{2} e |\psi_0(\hat{r}_B)|^2$.

The calculation of the force is a very straightforward matter and we will show how it is done for the symmetrical as well as antisymmetrical wave functions. In view of the fact that ψ is not precisely determined our solution is only approximate. It is

$$\psi_s e^{-2\pi i E_s t/\hbar} \quad \text{or} \quad \psi_a e^{-2\pi i E_a t/\hbar} \quad (26)$$

where E_s and E_a are energy functions for the two states both being functions of R . The force is given by $-dE_s/dR$ or $-dE_a/dR$.

It follows from (17) that

$$\int E |\psi|^2 d\tau = \int V |\psi|^2 d\tau - \frac{\hbar^2}{8\pi^2 m} \int \psi^\times \nabla^2 \psi d\tau.$$

In the rigorous solution E turns out to be when use is made of the normalization condition

$$E = \int V |\psi|^2 d\tau - \frac{\hbar^2}{8\pi^2 m} \int \psi^\times \nabla^2 \psi d\tau. \quad (27)$$

As we do not know ψ precisely we get the approximate values

$$E_s = \int V |\psi_s|^2 d\tau - \frac{\hbar^2}{8\pi^2 m} \int \psi_s^\times \nabla^2 \psi_s d\tau, \quad (28)$$

$$E_a = \int V |\psi_a|^2 d\tau - \frac{\hbar^2}{8\pi^2 m} \int \psi_a^\times \nabla^2 \psi_a d\tau. \quad (28')$$

The equations satisfied by $\psi_0(\hat{r}_A)$ and $\psi_0(\hat{r}_B)$ run as follows:

$$\nabla^2 \psi_0(\hat{r}_A) + \frac{8\pi^2 m}{\hbar^2} (E_0 - V_A) \psi_0(\hat{r}_A) = 0, \quad (29)$$

$$\nabla^2 \psi_0(\hat{r}_B) + \frac{8\pi^2 m}{\hbar^2} (E_0 - V_B) \psi_0(\hat{r}_B) = 0, \quad (29')$$

where

$$V_A = -e^2/r_A, \quad V_B = -e^2/r_B$$

and E_0 is the eigenvalue for the state ψ_{100} of hydrogen. Hence as ψ_s is real

$$E_s = \int \left(\frac{e^2}{R} + V_A + V_B \right) \psi_s^2 d\tau - \frac{\hbar^2}{8\pi^2 m} \int \psi_s \nabla^2 \psi_s d\tau \quad (30)$$

$$= E_0 + \frac{e^2}{R} + \frac{M + L}{1 + I}, \quad (30')$$

$$\text{where } M = \int V_A |\psi_0(\hat{r}_B)|^2 d\tau = \int V_B |\psi_0(\hat{r}_A)|^2 d\tau \quad (31)$$

$$\text{and } L = \int V_A \psi_0(\hat{r}_A) \psi_0(\hat{r}_B) d\tau = \int V_B \psi_0(\hat{r}_A) \psi_0(\hat{r}_B) d\tau. \quad (32)$$

Similarly,

$$E_a = E_0 + \frac{e^2}{R} + \frac{M - L}{1 - I}. \quad (30'')$$

The meaning of M is quite obvious. It represents the electrostatic energy of the 'part of the electron' at one nucleus in the field of the other. The exchange integral L represents a typical quantum effect. There is no classical analogue of it. It brings out the total effect of the interaction between the parts of the electron in the field of one of the nuclei. The values of L and M are as follows:

$$L = \frac{e^2 \alpha}{4} [1 + \alpha R] e^{-\alpha R}, \quad (32')$$

$$M = -\frac{e^2}{R} + \frac{2e^2}{R} (1 + \alpha R) e^{-2\alpha R}. \quad (31')$$

For evaluating L and M we make use of (20) and (21). The value of I is found out to be

$$I = \frac{1}{2} \left[1 + \alpha R + \frac{1}{2} (\alpha R)^2 \right] e^{-\alpha R}. \quad (33)$$

When substitutions of these values are made for I , L and M in (30') and (30'') we get the required functions E_s and E_a . It must be noted, however, that the values of E_s and E_a cannot be considered as reliable on account of the approximations adopted in the method. The nature of the $E_s - R$ and $E_a - R$ graphs is all that is significant. dE_s/dR is positive for some values of R . It becomes zero as R diminishes, becoming negative for smaller values of R . Hence there exists a critical distance for which the total energy E_s as a function of the parameter R is a minimum. The two atoms can come nearer under a mutual attraction but when the critical distance is reached there is a stable configuration. On the contrary, the behaviour of $E_a - R$ shows that there is no attraction between the atoms and that the smaller the value of R the greater is the force of repulsion.

Consider a hydrogen atom and a nucleus of the same element approaching each other. Suppose further that the relative velocity of the nuclei dies, on account of collision or radiation, when the two are near enough. If the electron happens to be in the antisymmetrical state ψ_a there would be a repulsion between the atom and the nucleus. If, however, the electron is in the state ψ_s a molecular ion would be formed. Our calculations are carried out for the ground state ψ_{100} . But we can similarly determine whether a stable ion can be formed when the electron in the field of either nucleus is originally in some excited state. The problem has been worked out by several investigators who have examined theoretically the stability of other states of H_2^+ .

In connection with the one-electron bond of the problem under discussion Frenkel² gives the analogy of the protoplasmic tie which connects two cells. When the probability distribution produces most of the charge cloud between the two nuclei the latter are at least partially shielded from the repulsion of each other; and it may happen that the force of attraction of the charge cloud triumphs over that of repulsion bringing the two nuclei nearer. This is what happens in the ψ_s -state considered above. In the other state the charge cloud is mostly distributed about the line AB produced both ways. This only increases the repulsion between the atom and the nucleus. The bonding electrons are responsible for the intermediate charge cloud, the antibonding ones going farthest away from the nuclei. Thus in the case of O_2 there are eight electrons from the two atoms participating most conspicuously in the readjustment of levels. Of these six are bonding and produce the

intermediate charge cloud responsible for the attraction. Two are non-bonding and their corresponding charge clouds are separated by the two oxygen nuclei. Usually the non-bonding electrons in pairs are bachelors with parallel spins. Thus the double bond of oxygen is explained by Herzberg's rule and the magnetic moment is accounted for by the two bachelor electrons.

We will apply the Heitler-London method to discuss the formation of the molecule H_2 . The two nuclei will be distinguished as before by the suffixes A and B and the two electrons by the suffixes 1 and 2. Thus (\hat{r}_{A_1}) is the vector from the nucleus A to the electron number one. Similarly $\psi(\hat{r}_{A_1})$ etc. are interpreted. Following (11) Heitler and London start with the symmetrical function

$$\psi_s = \frac{1}{\sqrt{\{2(1 + I^2)\}}} [\psi(\hat{r}_{A_1}) \psi(\hat{r}_{B_2}) + \psi(\hat{r}_{A_2}) \psi(\hat{r}_{B_1})] \quad (34)$$

where the condition of normalization is satisfied when we take

$$I = \int \psi(\hat{r}_{A_1}) \psi(\hat{r}_{B_1}) d\tau = \int \psi(\hat{r}_{A_2}) \psi(\hat{r}_{B_2}) d\tau. \quad (35)$$

It is assumed that $\psi(\hat{r}_{A_1})$ etc. are all normalized. It is definitely understood that both the electrons at A or B are in the same spatial level so that their spins are opposite. The very form of ψ_s implies that the resultant spin is zero. The corresponding antisymmetrical function which implies parallel spins is

$$\psi_a = \frac{1}{\sqrt{\{2(1 - I^2)\}}} [\psi(\hat{r}_{A_1}) \psi(\hat{r}_{B_2}) - \psi(\hat{r}_{A_2}) \psi(\hat{r}_{B_1})]. \quad (36)$$

Three distinct properties of ψ_s and ψ_a should be noted :

(1) For an interchange of the electrons, that is, for an interchange of the suffixes 1 and 2, ψ_s remains unaltered while ψ_a changes its sign. This is the property that we have in mind in calling one wave function symmetrical and the other antisymmetrical.

(2) If the nuclei are situated symmetrically about the origin on the axis of x , and if the x -coordinates of both the electrons change their signs ψ_s remains unchanged while ψ_a changes its sign. Thus ψ_s has the property of spatial symmetry with respect to the nuclei ; ψ_a has spatial antisymmetry in relation to them.

(3) An observable α which is symmetric in the spatial coordinates of the two electrons is always in the state ψ_s or in the state ψ_a provided only these two states are permissible. A transition from one state to another is not possible since

$$\int \psi_a^* \alpha \psi_s d\tau = 0, \quad (37)$$

identically, as the symmetry properties show.

The kinetic and potential energy of the system of two electrons in the field of two fixed hydrogen nuclei may be calculated straight from Schrödinger's equation or, from first principles, as shown below.

The kinetic energy of the electrons, on the classical theory, is

$$\begin{aligned} T &= \frac{1}{2} m (\dot{x}_1^2 + \dot{y}_1^2 + \dot{z}_1^2) + \frac{1}{2} m (\dot{x}_2^2 + \dot{y}_2^2 + \dot{z}_2^2) \\ &= \frac{1}{2m} (P_{x1}^2 + P_{y1}^2 + P_{z1}^2) + \frac{1}{2m} (P_{x2}^2 + P_{y2}^2 + P_{z2}^2). \end{aligned}$$

Hence the average kinetic energy in the state ψ_s is \bar{T} :

$$\begin{aligned} \bar{T} &= \int \psi_s^* T \psi_s d\tau \\ &= -\frac{\hbar^2}{8\pi^2 m} \int (\psi_s^* \nabla_1^2 \psi_s + \psi_s^* \nabla_2^2 \psi_s) d\tau \end{aligned} \quad (38)$$

where ∇_1^2, ∇_2^2 have the usual significance for the two sets of coordinates. Similarly, the average potential energy in the state ψ_s is \bar{V} :

$$\bar{V} = \int \psi_s^* \left(\frac{e^2}{R} + \frac{e^2}{r_{12}} - \frac{e^2}{r_{A1}} - \frac{e^2}{r_{A2}} - \frac{e^2}{r_{B1}} - \frac{e^2}{r_{B2}} \right) \psi_s d\tau. \quad (39)$$

These evaluations are carried out by making use of the equations which $\psi(\hat{r}_{A1})$ etc. satisfy, viz.,

$$\frac{\hbar^2}{8\pi^2 m} \nabla_1^2 \psi(\hat{r}_{A1}) + \left(E_0 - \frac{e^2}{r_{A1}} \right) \psi(\hat{r}_{A1}) = 0, \quad (40)$$

E_0 being the characteristic energy for $\psi(\hat{r}_{A1})$. For the other functions also E_0 is assumed to be the eigenvalue of energy. Let us denote $\bar{T} + \bar{V}$ by \bar{E} . There are two values of \bar{E} , corresponding to the wave functions, which we will distinguish by proper suffixes as \bar{E}_s and \bar{E}_a . $\bar{E}_s < 2E_0$, $\bar{E}_a > 2E_0$ within the degree of approximation for R . Hence $d\bar{E}_s/dR \geq 0$ in the domain of interatomic distances with the result that there is an attraction between the nuclei. If, however, the nuclei are brought nearer than a certain distance R_c attraction gives way to repulsion. The antisymmetric state does not favour chemical combination as the $\bar{E}_a - R$ graph shows. $2E_0 - \bar{E}_s$ is the heat of formation. This value cannot be taken seriously in view of the various approximations. In the first place, the very form (6) for ψ is inappropriate for quantitative discussion. For if (6) is true two electrons can come as near to each other as two non-interacting particles. But we know that each electron carries about it a space-pocket which is closed to another, on account of mutual repulsion. The exchange symmetry makes the approximate solution qualitatively valid; the spatial symmetry permits a collectivization of the two electrons by the nuclei. Hence the bond is particularly due to the latter.

It is futile to discuss the relative merits and demerits of the two methods just outlined. When more than two electrons are involved it is necessary to start with more general assumptions than those of either method. Even when there are only two valence electrons one may have to modify the method for dissimilar atoms participating in the chemical combination.

It is possible to explain directional valence by means of the wave functions of the valence electrons. When the valence electron is an s electron the wave function has spherical symmetry owing to $l = m = 0$ and no directional effect is witnessed. For a p electron there are three kinds of wave functions corresponding to the possibilities, $l = 1, m = 0, l = m = 1$ and $l = -m = 1$. They may be stated as follows

$$\begin{aligned}\psi_{n10} &= f_n(r) \cos \theta, \\ \psi_{n11} &= f_n(r) \sin \theta e^{i\phi/\sqrt{2}}, \\ \psi_{n1-1} &= f_n(r) \sin \theta e^{-i\phi/\sqrt{2}}.\end{aligned}$$

In the absence of an external field the wave functions of Schrödinger's theory are real and hence in the valence problems appear the wave functions $f_n(r) \sin \theta \cos \phi$ etc. The result of it is witnessed in the charge cloud which assumes a dumb-bell shaped form. For the molecular heat to be a maximum it is necessary that the exchange integral should be maximum. Hence the dumb-bell distributions of two electrons must overlap as much as wave-mechanical conditions permit. The directional effect of valence arises more or less from such overlapping of directed charge clouds.

3. SOME TYPICAL METHODS OF CALCULATIONS

In the more complicated calculations of valence and in the problem of n electrons a special mathematical property of the Schrödinger function is of great use. Let ψ be a normalized wave function so that

$$H\psi = E_0\psi. \quad (41)$$

In particular let E_0 be the value corresponding to the ground state of the system. Suppose further that ψ' is some other function which satisfies all the field and boundary conditions including that of normalization which ψ satisfies, except (41). Then it can be shown that

$$E' \equiv \int \psi'^* H \psi' d\tau > E_0. \quad (42)$$

For we can expand ψ' as follows:

$$\psi' = \sum_n a_n \phi_n \quad (43)$$

where ϕ_1, ϕ_2, \dots etc. are a complete set of normalized orthogonal functions each satisfying a Schrödinger equation of the form

$$H \phi_n = E_n \phi_n. \quad (44)$$

Hence $\psi'^* H \psi' = \sum_n a_n^* \phi_n^* H \sum_{n'} a_{n'} \phi_{n'}$

and $\int \psi'^* H \psi' d\tau = \sum_n a_n^* a_n E_n.$

Therefore

$$\begin{aligned} \int \psi'^* H \psi' d\tau - E_0 &= \sum_n a_n^* a_n E_n - E_0 \\ &= \sum_n a_n^* a_n (E_n - E_0) \end{aligned} \quad (45)$$

since by the condition of normalization it follows from (43) that

$$\sum_n a_n^* a_n = 1. \quad (46)$$

As E_0 is the least eigenvalue (42) follows from (45). When ψ is not precisely known one may try a number of plausible functions ψ' in order to get at an estimate of E_0 . What is most helpful in these calculations is that even when ψ' differs from ψ appreciably the corresponding deviation of E' from E_0 is, in most cases, rather small. The method has been successfully employed to determine the value of energy in the ground state of the normal helium atom. When the eigenstates are not discrete but continuous an integral replaces the right-hand side of (43). The procedure is substantially unaltered for this case. The variational method is known also as the Ritz method and it is most helpful in practice for calculating the molecular heats.

In valence problems and in several theoretical investigations of polyatomic molecules the group theory is of great service. Its great advantage is that some of the inner structural relations of mathematics are disclosed as fundamental properties of a group without having to solve intricate differential equations and plough through devious approximations. Thus the selection and polarization rules for polyatomic molecules can be deduced either by quantum mechanical methods or by the application of the group theory to the classical treatment. It is true of course that there is a limitation on the results that can be obtained purely by the group theory. To know how the group methods are applied and how certain interesting relations are obtained therefrom the reader may find it worthwhile to take up certain very elementary problems first.

Suppose that we want to know the number of persons present at a meeting. Somebody counts and tells that the number is, say, 45. You question the accuracy of the calculation and then the counting is done openly in order to convince you. Even if the counting yields the same result you would be justified in not accepting the result provided you do not know fundamental properties of cardinals and ordinals. Actually what happens in most cases is just the reverse: you do not know the properties and are not even aware of their existence and hence you usually accept the result of a count on account of convention and experience. But in doing this you are either illogical or unconsciously making use of the theory of groups. If a particular count gives the result, 45, what is the guarantee that a different count would give the same result? Unless that guarantee is first secured it is not right to rely upon the result of a particular count. Moreover, if different counts are

Thus A_1 is I and $A_1 A_2 = A_2$. Similarly $A_3 A_4$ means first substituting a, c, b for a, b, c and then substituting c, a, b for a, c, b . The result is equivalent to the substitution cab for a, b, c , that is, A^6 . So $A_3 A_4 = A_6$. This group is called symmetric for the reason that its operations leave any symmetric function of a, b, c unchanged in value.

A select set of operations of the group may by themselves satisfy all the postulates of a group and thus constitute a sub-group. Take any element A of a group. Then A^2, A^3, \dots etc. must be also elements of the group. It is certain that $A, A^2, \dots A^r$, etc. cannot have more than n distinct values since n is the order of the parent group of which A is a member. Let $A^r = I$. Then it is clear that the higher powers of A will not give any new elements. Hence the powers of A constitute a sub-group of order r and it is called a cyclic subgroup.

When the elements of two groups have a one-to-one correspondence they are said to be isomorphic. Thus if A corresponds to A' and B to B' , AB must correspond to $A'B'$. Unless the last condition is fulfilled the groups would not be isomorphic. Further, if one of the groups represents analytical operations it is said to give a representation of the other group which consists of purely abstract operations. Such a representation of the cyclic group, ($A^r = I$), is given by the elements, $e^{2\pi i p/r}$ where $p = 0, \dots (r-1)$.

Let $\{G\}$ be a group of order n and let $\{H\}$ be a sub-group of order m . It is obvious that n/m must be an integer. For consider an element A_1 which does not belong to $\{H\}$. Then HA_1 are m new elements which do not belong to $\{H\}$. If this is not true either $H_p A_1 = H_q$, that is $A_1 = H_p^{-1} H_q$ or $H_p A_1 = H_q A_1$ that is, $H_p = H_q$ both of which are impossible if H_p and H_q belong to $\{H\}$ and A_1 does not. Let A_2 be an element of $\{G\}$ which does not belong to $\{H\}$ and also to the set HA_1 . Then HA_2 is another set of elements which can be similarly shown to be new. This process must come to a stop somewhere say at A_{p-1} . Then all the elements of $\{G\}$ are the mp elements, H, HA_1, \dots, HA_{p-1} . Hence $n = mp$. Also it can be shown that, if $\{H\}$ is a self-conjugate group

$$H, HA_1, HA_2, \dots, HA_{p-1} \quad (47)$$

is itself a group of order p . It is usually denoted by $\{G/H\}$. What is meant by conjugate operations will now be explained.

Let A, B, S be any three elements of $\{G\}$. If

$$B = SAS^{-1} \quad (48)$$

A and B are said to be conjugate to one another. If A and B are identical $AS = SA$. An important property of conjugate operations is that, if $A^r = I$ then

$$B^r = (SAS^{-1})^r = SA^r S^{-1} = I. \quad (49)$$

Now all the elements which commute with any element A form a group. For if S and T are two such elements of $\{G\}$

$$S = ASA^{-1}, \quad T = ATA^{-1}$$

and hence $ST = ASA^{-1}ATA^{-1} = ASTA^{-1}$.

Let us call this group $\{H\}$. Then using the result (47) we know that

$$A, A_1^{-1} A A_1, A_2^{-1} A A_2, \dots, A_{p-1}^{-1} A A_{p-1} \quad (50)$$

are all the conjugate elements of which A is one. They constitute what is called a class and its order is an integer. The elements of the group can be divided into such classes.

Consider now a set of substitutions

$$\chi_r = \sum \alpha_{rs} \chi'_s \quad (51)$$

which is also expressed as

$$\chi = A\chi' \quad (52)$$

where χ and χ' are m -dimensional vectors with components χ_r and χ'_s and A is the matrix $|\alpha_{rs}|$. If

$$\chi' = B\chi'' \quad , \quad \chi'' = C\chi''' \quad (52)$$

are similar substitutions

$$\chi' = AB\chi''' \text{ and } \chi = ABC\chi'''.$$

Here A, B, C , etc. constitute a group. Every finite group is isomorphic with some group of linear substitutions. Hence matrices can always be associated with the abstract operations A, B, C etc. Suppose that by choosing a suitable transformation all the matrices of the group can be made to appear as

$$A = \begin{pmatrix} A_1 & 0 & 0 & \dots & 0 \\ 0 & A_2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & A_p \end{pmatrix} \quad (53)$$

where A_1, A_2, A_p are square matrices. In general when the matrix A_r is repeated α_r times in such a representation we put

$$A = \sum_{r=1}^p \alpha_r A_r \quad (54)$$

The sum of the diagonal elements of A is called the character and is shown as $\chi(A)$. This has great importance in applications since if $B = SAS^{-1}$,

$$\begin{aligned} \chi(BS) &= \chi(SA) \\ \chi(B) &= \chi(SAS^{-1}). \end{aligned} \quad (55)$$

and hence

If A denotes the matrix of an observable and B is its diagonalized form we know that the diagonal elements of B give the eigenvalues. The sum of the diagonal elements has therefore a physical significance. The orthogonal property of wave functions is essentially the result of the orthogonality of group characters. The following two properties of the group characters are very helpful in applications.

(1) If the group can be split up into p different classes there are p distinct parts into which A can be split up, A_1, A_2, \dots etc. so that none of these can be reduced further into the form

$$\begin{pmatrix} P_1 & O \\ O & P_2 \end{pmatrix}$$

where P_1 and P_2 are square matrices. The group characters for the parts satisfy the condition,

$$\chi(A) = \sum_{j=1}^p \alpha_j \chi(A_j). \quad (56)$$

(2) The characters are subject to the relations

$$\sum \chi(A_j) \chi(A_{j'}) = h \delta_{jj'} \quad (57)$$

$$\text{and} \quad \sum_{j=1}^p [\chi(A_j)]^2 = h. \quad (58)$$

Here we have h as the order of the group and the summation in (57) is made for all the operations of the group. In quantum mechanical studies of the spectra properties of the symmetric group are extensively illustrated.

One property which we wish to mention last concerns direct products. Suppose we have two vectors X and Y of dimensions m and n and with matrices of transformation A, B :

$$X = AX', \quad Y = BY'$$

$$\text{or} \quad x_i = \sum a_{i\alpha} x'_\alpha, \quad y^j = \sum b_{j\beta} y'_\beta.$$

The direct product of X and Y provides a vector Z of dimensions mn and the law of transformation is

$$Z = CZ' \text{ or } z_i y_j = \sum \sum a_{i\alpha} b_{j\beta} x'_\alpha y'_\beta. \quad (60)$$

Here the mn components are $x_i y_j$, $i = 1, 2, \dots, m, j = 1, 2, \dots, n$. Hence

$$\chi(C) = \chi \left| a_{i\alpha} b_{j\beta} \right| = \sum a_{i\alpha} \sum b_{j\beta} = \chi(A) \chi(B). \quad (61)$$

From any transformation of the form

$$\begin{aligned} x' &= ax + by \\ y' &= cx + dy \end{aligned}$$

it is possible to determine the law of transformation for a vector whose components are

$$x, \quad x^{f-1}, \quad y, \dots, x^{f-r}, \quad y^r, \dots, y^f.$$

Thus the laws of transformation connecting the $(f+1)$ -dimensional space become known. Similarly we can consider the laws of transformation for antisymmetric functions of the form

$$x_i y_k - x_k y_i$$

when x_i and y_i are vectors of n dimensions. The matrices of transformation of antisymmetric tensor components can also be multiplied as in (60) and the result can be expressed in the reduced form as in (53). This gives the Clebsch-Gordon series.

Suppose³ that there are two atoms whose outermost subshells contain electrons f_1 and f_2 . We consider the two nuclei as fixed at some distance apart. The interaction energy W has a definite symmetry for both the atoms and its laws of transformation give us a matrix $A' = A_1 \times A_2$. The matrix is invariant for all symmetrix transformations. This is expanded in the vector space of electron spin and the irreducible components give us the valence bonds and the states of symmetry of the molecule. The procedure can be adopted for more complicated chemical combinations. The numbers of external electrons corresponding to the states in the expansion, when compared to f_1 and f_2 tell us what electrons are bonding or antibonding.

One of the most important applications of groups is in the study of crystallography. Each crystal is characterized by an expression of elastic energy quadratic in the six strains and involving twenty-one constants. There are thirty-two different groups some one of which leaves the energy expression invariant for any crystal. The nature of symmetry in the crystal determines the groups which provide relations between the constants.

4. OTHER FORCES OF QUANTUM-MECHANICAL NATURE

The subject that comes under the title of this section is very vast. First there are the forces of cohesion which explain the surface tension phenomena and latent heats and as they are responsible for the constant a of van der Waals' equation,

$$\left(p + \frac{a}{v^2} \right) (v - b) = RT,$$

they are also referred to as van der Waals' forces. For smaller distances the forces of repulsion between atoms are more powerful. They account for the rigidity of matter and also for the other elastic properties. When the atoms are brought still nearer together so that there is a partial overlapping of the electron clouds, again the forces become attractive and formations of

molecules are witnessed. At still smaller distances the forces of repulsion between the electrons arrest our attention. The Fermi-Dirac statistics of antisymmetrical states describes the full effect of interactions in large assemblies of electrons. Once more and for intra-nuclear distances forces of attraction assert themselves. There is evidence that two protons attract each other within a nucleus. The nuclear forces are treated as exchanges forces and we will give some account of them. It is not possible to touch here the topic of the constants of elasticity as that would take us too far into a discussion of metals and crystal structure. The orthodox treatment⁵ of the van der Waals forces between two molecules was based on the electrostatic forces given by the usual expansion for the potential of a charge assembly,

$$V = \frac{\sum e_i}{R} + \sum_{\rho=1}^3 P_{\rho} \frac{\partial}{\partial x_{\rho}} \left(\frac{1}{R} \right) + \frac{1}{2!} \sum \Theta_{\rho\sigma} \frac{\partial}{\partial x_{\rho}} \frac{\partial}{\partial x_{\sigma}} \left(\frac{1}{R} \right) + \dots \quad (62)$$

where $P_1 = \sum e_i x_i, \quad P_2 = \sum e_i y_i, \quad P_3 = \sum e_i z_i$
 $\Theta_{11} = \sum e_i x_i^2, \quad \Theta_{12} = \sum e_i x_i y_i, \text{ etc.}$
 $R^2 = X_1^2 + X_2^2 + X_3^2;$

and X_1, X_2, X_3 are the coordinates of the current point, the origin being at the centre of gravity of the charges. For a neutral molecule $\sum e_i = 0$ and the first term in the expansion (62) vanishes. At great distances the field of the molecule may therefore be that of a dipole if the second term is non-zero or that of a quadrupole if that term also vanishes. If the third term is also zero the molecule behaves as an octupole. Equal and opposite charges separated by a small distance give us a dipole. Alternately equal and opposite charges at the corners of a parallelogram have the field of a quadrupole. But if the charges on all the corners of a parallelepiped are so distributed that they are alternately equal and opposite the field of an octupole is obtained. The degree of polarity varies from molecule to molecule. We will give here the explanation of intramolecular forces first given by Heitler and London on the assumption that the charge distribution of molecules has the field of a harmonic electrical oscillator. They were guided in this by the consideration that the energy values of an oscillator account for the energy of molecules at the absolute zero. For two oscillators⁶ separated by a distance R and in the same line the Hamiltonian is

$$H = \frac{1}{2m} (P_1^2 + P_2^2) + \frac{e^2}{2a} (x_1^2 + x_2^2) + 2 \frac{e^2 x_1 x_2}{R^3} \quad (63)$$

when higher terms in the expansion of the interaction energy,

$$\frac{e^2}{R} + \frac{e^2}{R + x_1 + x_2} - \frac{e^2}{R + x_1} - \frac{e^2}{R + x_2} \quad (64)$$

are neglected. If the last term of (63) is absent the oscillators are uncoupled and the frequency is

$$\nu_0 = \frac{1}{2\pi} \sqrt{\frac{e^2}{am}} \quad (65)$$

The constant a in (63) and (65) becomes known when the law of the restoring force in the oscillators is given. It is a measure of the distortion undergone by the molecules. Using the transformation

$$x_s = \frac{1}{\sqrt{2}} (x_1 + x_2), \quad x_a = \frac{1}{\sqrt{2}} (x_1 - x_2)$$

we get $\frac{\partial H}{\partial \dot{x}_s} = p_s = \frac{1}{\sqrt{2}} (p_1 + p_2), \quad \frac{\partial H}{\partial \dot{x}_a} = \frac{1}{\sqrt{2}} (p_1 - p_2) = p_a.$

Hence

$$H = \left\{ \frac{1}{2m} p_s^2 + \left(\frac{e^2}{2a} + \frac{e^2}{R^3} \right) x_s^2 \right\} + \left\{ \frac{p_a^2}{2m} + \left(\frac{e^2}{2a} - \frac{e^2}{R^3} \right) x_a^2 \right\} \quad (66)$$

This gives us two oscillators with frequencies

$$\nu_s = \frac{1}{2\pi} \left[\frac{e^2}{m} \left(\frac{1}{a} + \frac{2}{R^3} \right) \right]^{\frac{1}{2}}, \quad \nu_a = \frac{1}{2\pi} \left[\frac{e^2}{m} \left(\frac{1}{a} - \frac{2}{R^3} \right) \right]^{\frac{1}{2}}, \quad (67)$$

and hence the total energy, when the quantum laws are obeyed, is

$$h \nu_s (n_s + \frac{1}{2}) + h \nu_a (n_a + \frac{1}{2}) \quad (68)$$

where n_s and n_a are two integers. In the ground state, $n_s = 0$, $n_a = 0$ and

$$E_0 = \frac{1}{2} h (\nu_s + \nu_a) = h \nu_0 \left(1 - \frac{a''}{2K^3} \right). \quad (69)$$

The force between the molecules is $-dE/dR$ when they are in the state E . The force is one of attraction and varies as R^{-7} . Thus even in the absence of chemical forces neutral molecules with valencies saturated may be attracted to each other at low temperatures. Illustrations are furnished by metallic molecules such as Hg_2 .

Various¹⁰ theories of the neutron-proton interaction in a nucleus have been offered by Heisenberg, Majorana, Wigner and Ukawa. As the method adopted by Ukawa suggests the coming into existence of particles such as the heavy electrons observed in the cosmic ray phenomena we will briefly describe his procedure. It must be pointed out, however, that Eddington has argued from the standpoint of relativity to show that the particles of the Ukawa theory are just comparison particles which cannot be identified with any real particles such as mesotrons or the heavy electrons. Eddington's argument is very difficult to follow.

Ukawa's theory originates from the classical equation of electrodynamics,

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \phi = 4\pi\rho \quad (70)$$

ρ being the electric density. The modified equation,

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 + \lambda^2 \right) \phi = 4\pi g\rho \quad (71)$$

is taken, at the start, involving the two constants λ and g . The solution of the new equation gives the nuclear potential

$$\phi = \frac{1}{4\pi} g \frac{e^{-\lambda r}}{r}, \quad (72)$$

at a distance r from a proton or a neutron. It is obvious that ϕ falls off rapidly as r increases; this is as it should be for an observed characteristic of all nuclear forces is their short range. Replacing the differential operators by the energy and momentum symbols in (71) we have

$$\left(-\frac{E^2}{c^2} + p_x^2 + p_y^2 + p_z^2 + \lambda^2 \frac{h^2}{4\pi^2} \right) \phi = 4\pi g\rho \frac{h^2}{4\pi^2}. \quad (73)$$

In view of the smallness of the right-hand side term, (73) may be taken as equivalent to

$$\left(-\frac{E^2}{c^2} + p_x^2 + p_y^2 + p_z^2 + \lambda^2 \frac{h^2}{4\pi^2} \right) \phi = 0 \quad (74)$$

But this is just the wave equation for a particle having energy E , momentum (p_x, p_y, p_z) and mass M :

$$M = \frac{\lambda h}{2\pi c}.$$

From (72) one concludes that $r = \frac{1}{\lambda}$ may be taken as a good measure of the range of the nuclear forces. The range is known to be of the order of e^2/mc^2 . This gives an estimate of λ . The mass of the particle is of the order of $250m$. The emission and absorption of such a particle is the cause of nuclear interaction. The constant g in (72) takes account of the nuclear field.

The heavy particle which may be positive or negative or even neutral (neutretto) is known to be unstable. The heavy electron rapidly decays with the emission of radiation into a β particle. The conservation of angular momentum now necessitates the acceptance of the neutrino postulate. For as the neutron, the proton and the electron have each a spin $\pm\frac{1}{2}$, the transition, neutron \rightleftharpoons proton, cannot take place with the emission or absorption of a heavy electron unless another particle, the neutrino, with a spin $\pm\frac{1}{2}$ is induced to participate in the phenomenon.

April 15, 1940.

[Addendum :—Conscious of many imperfections and lacunæ I conclude the Report of my Lectures at this stage. I should have liked to include a more thorough treatment of the group theory and its applications and also of the nuclear phenomena. All that must await another occasion to see the light of day.—V. V. N. (Dec. 20, 1941)]

REFERENCES

- (1) VAN VLEEK and A. SHERMAN, "*Rev. Mod. Phys.*" 7, 3, pp. 168-228 (1935).
- (2) FRENKEL, J., "*Wave Mechanics*", Part I, pp. 279-309 (1936).
- (3) WEYL, H., "*The Theory of Groups and Quantum Mechanics*", p. 371 (1939).
- (4) ROSENTHAL and MURPHY, "*Rev. Mod. Phys.*" 8, 4, pp. 317-46 (1936).
- (5) BETHE and BACHER, "*Rev. Mod. Phys.*" 8, 2, pp. 92-107 (1936).
- (6) BORN, M., "*Atomic Physics*" pp. 253, 340-2 (1935).
- (7) MILLER, BLICHFELDT and DICKSON, "*Finite Groups*", (1938).
- (8) MOTT, N. F. "*An outline of Wave Mechanics*", p. 91 (1930).
- (9) PAULING and WILSON, "*Introduction to Quantum Mechanics*", pp. 180-4 (1935).
- (10) BHABHA, H. J., "*Bombay Univ. J.*" 8, 3, pp. 34-6 (1939).

— THE END —

NAME INDEX

- Bacher, 156.
 Beams, 26.
 Bethe, 117, 156.
 Bhabha, 156.
 Birtwistle, 23, 45, 87.
 Blichfeldt, 156.
 Bohr, 31, 34, 73, 116, 134.
 Born, 45, 56, 59, 71, 106, 156.
 Bowen, 117.
 Brillouin, 71.
 Burges, 34.

 Chadwick, 117.
 Compton, 104.
 Condon, 22, 45, 70, 71, 106, 117, 134.
 Curie, 116.

 De Broglie, 29, 45, 48, 52, 71, 101.
 Democritus, 24.
 Dickson, 156.
 Dirac, 12, 23, 45, 75, 76, 85, 88, 106, 138.

 Eckart, 83.
 Eddington, 115, 121, 134, 154.
 Ehrenfest, 63.
 Einstein, 24, 27, 38, 41, 44.
 Eisenhart, 23.

 Fermi, 115, 116.
 Flint, 106, 115.
 Franck, 35, 39.
 Frenkel, 23, 31, 45, 71, 106, 109, 117, 138, 143, 156.

 Gamow, 61.
 Goudsmit, 104.

 Hahn, 116.
 Hartree, 22.
 Heisenberg, 12, 45, 71, 87, 106, 138, 154.
 Heitler, 42, 45, 131, 134, 137, 144.
 Hertz, 35, 39.

 Herzberg, 71, 103, 114, 117, 136, 144.

 Jeans, 23, 115.

 Kramers, 72.
 Kronecker, 26.

 Lawrence, 26, 117.
 Lennard-Jones, 138, 139.
 Lewis, 137-39.
 Lienard, 119.
 London, 137-39, 144.
 Lorentz, 28.
 Lucretius, 24.

 Mach, 3, 22.
 Majorana, 154.
 Margenan, 134.
 Maxwell, 23.
 Meitner, 116.
 Mendeléeff, 33.
 Miller, 156.
 Millikan, 134.
 Milne, 31.
 Mises, 22.
 Morse, 45, 71, 106.
 Moseley, 33.
 Mott, 71, 156.
 Mullikan, 139.
 Murphy, 156.

 Narlikar, 117.
 Noddack, 116.

 Oppenheimer, 61.
 Osgood, 22, 23.

 Pauling, 71, 156.
 Pearson, 22.
 Planck, 24, 25, 30.
 Podlosky, 51.

 Ramsauer, 61.
 Reiche, 31, 56.

- Richardson, 115.
Ritz, 147.
Rosenthal, 156.
Ruark, 23, 45, 85, 27.
Rupp, 52.
Russell, 22.
Rutherford, 31, 56, 117.
- Savitch, 116.
Schrödinger, 22, 34, 42, 45, 49, 63, 83.
Sherman, 156.
Shortley, 22, 70, 117, 134.
Sommerfeld, 36, 40, 67.
- Tait, 1, 22.
Thomas, 109.
Thomas, L., 68, 104.
Thomson, G. P., 52.
Thomson, J. J., 31.
- Thomson, W., 1, 22.
Tolman, 23, 106.
Turner, 117.
- Uhlenbeck, 104.
Ukawa, 154.
Unsöld, 70.
Urey, 23, 45, 85, 87.
- Van Vleck, 156.
Viechart, 119.
- Watson, 134.
Weyl, 22, 156.
Wheeler, 116.
Whittaker, E. T., 23.
Wigner, 154.
Wilson, 71, 156.
- Zimmer, 71.
-

SUBJECT INDEX

- Adiabatic invariants, 33, 34.
- Angular momentum, 67, 103.
- Auger effect, 33.
- Average velocity, 4, 100.
- Avogadro's number, 27, 28.

- Balmer series 36.
- Black-body radiation, 24, 30.
- Bohr's theory, 16, 17, 31, 51, 53, 59, 67; postulates and hypotheses, 32; successes, 35, 40; failures, 43-45, 55; frequency, 34, 72; quantum numbers, 52, 68; energy, 66; orbit, 69; line breadth, 125.
- Bremsstrahlung, 129.

- Canonical equations, 8.
- Causation, 1, 59.
- Classical statistics, 42.
- Collectivization, 139-46.
- Collision experiments, 39, 60-63.
- Compton effect, 40, 41, 43; shift, 41.
- Conditionally periodic system, 16, 72.
- Conservation principles, 134.
- Contact transformation, 8-9, 15, 78. •
- Correspondence principle, 32, 34, 37, 43, 73.
- Coupling, Russell-Saunders and otherwise, 114.

- D'Alembert's principle, 5.
- Dalton's hypothesis, 25.
- De Broglie waves, 50, 61, 68.
- Degenerate system, 16, 19, 20, 32, 38, 78, 80.
- Dipole moment, 69.
- Dirac's theory, 34, 58, 59, 91, 94; quantum conditions, 14; electron spin, 68; non-commutative relations, 77, 90; postulates and principles, 95-100; p - q transformation, 101; relativistic invariance, 104-106, 127; hole theory, 132-33.
- Distinguishability, 3, 42, 107.

- Doppler effect, 127.
- Dulong and Petit's law, 27.

- Ehrenfest's adiabatic principle, 17, 67, 113.
- Eigen- ϕ , 55.
- Eigenvalue, 55.
- Einstein's hypothesis, 26, 28; relations, 29, 40, 101; failures of the theory, 42-43.
- Einstein-Bose statistics, 43, 110, 111.
- Election: spin, 43, 103-106; bonding, non-bonding, anti-bonding, 136; 'vagabonding', 138; bachelor, 114, 137; heavy, 154.
- Exchange relations, 15.
- Expectation values, 99, 102.

- Faraday's laws, 24.
- Fermat's principle, 46-48.
- Fine-structure constant, 37, 67.
- Forbidden transitions, 32, 33, 69, 114.
- Force, 2, 135.

- Green's theorem, 54, 58.
- Group theory, 91, 147-52.
- Group velocity, 49, 50, 57.

- Hamiltonian, 13.
- Hamilton-Jacobi equation, 9, 15.
- Hamilton's principle, 7.
- Harmonic oscillator, 18, 25, 52, 80-83, 119, 122, 153.
- Hermite polynomial, 53.
- Hermitian space, 93, 96.
- Herzberg's rule, 136, 144.
- Heisenberg's uncertainty principle, 56, 100-103, 124, 130; matrix theory, 72-84.
- Hund's rule, 114.

- 'Inside the electron', 4, 129.
- Interpretation of ϕ , 54-59, 68, 141.

- Kepler problem, 64.
 Kronecker delta, 86.
 Lagrangian, 7.
 Langer polynomials, 66, 68.
 Landé g -factor, 45, 104.
 Law of equipartition, 28.
 Law of gravitation, 2.
 Laws of motion, 1-2, 63.
 Least action, principle of, 7.
 Legendre functions, associated, 65.
 Light, nature of, 123.
 Lorentz transformation, 29, 30, 104,
 105, 120, 121.
 Lyman series, 36.
 Maxwellian distribution, 24.
 Maxwell's field equations, 120, 133.
 'Microscopic' state, 4.
 Metastable state, 114.
 Multiple processes, 128.
 Multiplicity of term, 112.
 Natural line breadth, 126.
 Neutrino, 155.
 Non-commutative relations, 12, 14, 51,
 77, 95, 122.
 Non-compatible observables, 90, 100.
 Nuclear disintegration, 70; fission, 116.
 Observables, 72-73, 91.
 Operators, 91; differential, 50, 51, 75.
 Paschen-Back effect, 104, 114.
 Pauli's exclusion principle, 64, 110, 111,
 132, 137.
 Periodic table, 33, 40, 64, 111.
 Permutation operator, 107.
 Pfaffian, 76.
 Phase velocity, 46.
 Photo-electric effect, 41, 43.
 Photon, 21, 42, 128, 131.
 Planck's constant, 14, 29; hypothesis, 18,
 26, 28, 34; radiation formula, 25, 27,
 42.
 Poisson brackets, 12-15, 75, 77, 96.
 Polarization, 37, 43, 70, 120.
 Positron, 131.
 Probability, 56, 59, 85-87, 95, 98.
 Quadrupole moment, 69.
 Quantum conditions, 14, 32, 52, 76, 96.
 Quantum numbers, 34, 52.
 Raman effect, 41.
 Rayleigh's radiation formula, 24, 25.
 Relativistic Hamiltonian, 36.
 Ritz's method, 147.
 Rutherford's model, 31.
 Rydberg constant, 31, 36.
 Rydberg-Ritz combination principle, 73.
 Schrödinger's wave function, 31, 91, 100,
 110, 146; equation, 46, 50, 64, 67, 94,
 95, 108, 145; postulates, 50-52;
 theory, 72, 80, 84.
 Shells, 66.
 Stark effect, 31, 35, 38.
 Stephen's law, 25.
 Superposition, principle of, 96.
 Symmetry properties, 108.
 Transverse self-energy, 54, 130.
 Valence, 135; directional, 146.
 Van der Waals forces, 53, 152.
 Wave velocity, 46, 50.
 Wien's formula, 24.
 Wilson-Sommerfeld conditions, 32, 51.
 Zeeman effect, 45, 104.
 Zeno's paradox, 4, 56.
 Zero-point energy, 53.

